

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:17:37 ; Search time 37.71 Seconds
(without alignments)
3550.731 Million cell updates/sec

Title: US-09-803-165-34
Perfect score: 4026
Sequence: 1 MRPDYLIRKDGKPIRIRK.....KEDLKYSQKQGLDAMLKK 774

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_unclassified:*
14: SP_rviro:*
15: SP_rviro:*
16: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3212	79.8	775	1 09P9K4	09P9K4 pyrococcus
2	3211	79.8	775	1 09H06	09H06 pyrococcus
3	3081	76.5	759	1 09H98	09H98 pyrococcus
4	1204.5	29.9	781	1 09P9N1	09P9N1 sulfurispha
5	1204.5	29.9	781	1 09P9Y1	09P9Y1 sulfobolus
6	1190	29.6	803	1 059691	059691 pyrococcus
7	1125	27.9	785	1 09P9M2	09P9M2 pyrococcus
8	774	19.2	195	1 09H85	09H85 thermococcus
9	631.5	15.7	1443	5 09NAH1	09NAH1 caenorhabdi
10	626	15.5	901	17 09HRV9	09HRV9 halobacteri
11	623	15.5	796	17 09HJRO	09HJRO thermoplasma
12	622.5	15.5	800	17 097AH3	097AH3 thermoplasma
13	616.5	15.3	876	17 0971C7	0971C7 sulfobolus
14	616.5	15.3	1105	11 091VT0	091VT0 mus musculus
15	616	15.3	1107	4 096H98	096H98 homo sapien
16	584.5	14.5	914	1 059690	059690 pyrococcus

17	567	14.1	1492	10 09FHA3	09FHA3 arabidopsis
18	559.5	13.9	223	17 026310	026310 methanother
19	555.5	13.8	1458	5 09D56	09D56 xenopus lae
20	527	13.1	1472	5 09D50	09D50 drosophila
21	516.5	12.8	844	1 031096	031096 cenarchaeum
22	509	12.6	845	1 074046	074046 cenarchaeum
23	507.5	12.6	1001	12 090626	090626 porcine lym
24	506.5	12.6	1004	12 090628	090628 porcine lym
25	506.5	12.6	1415	5 077034	077034 drosophila
26	504.5	12.5	2154	10 092VC8	092VC8 arabidopsis
27	491	12.2	913	12 092Z29	092Z29 chlorella v
28	488	12.2	2207	3 093845	093845 porcine cyt
29	487	12.1	1007	12 091F13	091F13 porcine cyt
30	486	12.1	1007	12 091F15	091F15 porcine cyt
31	486	12.1	1007	12 091F16	091F16 porcine cyt
32	484	12.0	1855	5 093X75	093X75 plasmodium
33	484	12.0	1855	5 09BHNO	09BHNO plasmodium
34	483	12.0	2271	10 09SGD5	09SGD5 arabidopsis
35	481	11.9	1007	12 091F17	091F17 porcine cyt
36	481	11.9	1009	12 09YMQ4	09YMQ4 ateline her
37	481	11.9	1912	5 09UOH1	09UOH1 plasmodium
38	479	11.9	1016	12 0993K6	0993K6 callitrichi
39	478.5	11.9	1013	12 092827	092827 human herpe
40	477.5	11.9	774	2 09F175	09F175 pseudomonas
41	477	11.8	2243	4 09UNF3	09UNF3 homo sapien
42	477	11.8	2286	4 09Y554	09Y554 homo sapien
43	477	11.8	2297	4 09Y555	09Y555 homo sapien
44	471.5	11.7	1041	12 09DKT8	09DKT8 elephant he
45	468.5	11.6	1012	12 040910	040910 kaposi's sa

ALIGNMENTS

RESULT 1
ID 09P9K4 PRELIMINARY: PRT: 775 AA.
AC 09P9K4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7).
GN DNA POL.
OS Pyrococcus glycoformans.
CC Archaeae; Euryarchaeota; Thermococcales; Pyrococcus.
OX NCBI_TaxID=74610;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AL585;
RA Querellou J., Cambon M.A., Lesongeur F., Forterre P., Barbier G.;
RT "DNA polymerase genes organisation of species belonging to
Thermococcales and phylogenetic implications."
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N
PYROPHOSPHATE + DNA(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
EMBL: AJ004834; CAB81809.1; -
DR HSSP: P56689; JNCO.
DR InterPro: IPR002064; DNA_POL_B.
DR Pfam: PF00136; DNA_POL_B; 1.
DR Pfam: PF03104; DNA_POL_B_exo; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLB; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN.1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase;
Nucleic acid transferase; Transferase.
SQ SEQUENCE 775 AA; 90492 MW; 677264920D770C0F CRC64;

Query Match 79.8%; Score 3212; DB 1; Length 775;
Best Local Similarity 77.1%; Pred. No. 1.1e-180;
Matches 595; Conservative 86; Mismatches 89; Indels 2; Gaps 1;

```

QY 1 MFDYDITKDGKPIIRIFKENGEEFKIELDPHPQPIYALALDKSDAIDEIKAKGERHG 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MILDADYITDEGKPIIRIFKENGEEFKVEYDRNFRPIYALLDKDSDQIDEVKKITAEHRG 60
QY 61 KIYRVDAVAVKKKFLGRDVEVKKLFEHPQDYPALRGKIREHPAVIDIYEDIPPAKRY 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 KIYRIVDVEVKKKFLGRPLEVKKLFEHPQDYPALRGKIREHPAVAVDIYEDIPPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEFKGETIMISYADEEARVITKNIDLPY 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEFKGETIMISYADEEGAKVITKKKADLPY 180
QY 181 VDVVSNEREKIKRFVQIVREKDPDLITVNGDNFDPYLKRAEKGVLTLGRDKEHPE 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 VEVVSSEREKIKRFVQIVREKDPDLITVNGDSFDPYLKRAEKGVLTLGRDGS--E 238
QY 241 PKIHRGDSFAVEIKGRHFDLPVVRRTINLPTYLEAVEYALCKTSKGLAEETAI 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 PKMORLGDMTAVEIKGRHFDLVHVRRTINLPTYLEAVEYALFGPKKAVAHETAE 298
QY 301 WETEESMKLAQYSMEDARATYELGKEFFPMEAEALKIGOSVADYSRSSTGNLVEYLL 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 WETGKGLERAKYSMDAKVTYELGREFPMEANQSLVGOPLMDYSRSSTGNLVEYLL 358
QY 361 RAVARNELAPNKPDEBEYRRRLRTTYLGIVKEPERGLMENTIYDFRCLYSIIVTN 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 RKAVERNELAPNKPDEREYERRLRRESYAGGYVEPEKGLMEGLVSDFRSLYPSIIITHN 418
QY 421 VSPDTLEREGCKNVDVAPYIGYKFCDFPFTISIGELITMROEIKKKKATIDPEKK 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 VSPDTLNRBECMEYDVAPEVKKHFCDFPFTISLKLRLDEROEIKRRKAKSDPLEKK 478
QY 481 MLDYRORAVKLANSYGYGYGPKARWYKSCAESYATAMGRHYIEMTIKEIEEKFGEKVL 540
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 MLDYRORAIKILANSYGYGYGAKARWYKSCAESYATAMGREYIEFVRKELEKFGFKVL 538
QY 541 YADTDFYATIPGKPEPTIKKAKKEFLKYINSLPGILLEVEYGFYLRGFFVAKKRAYI 600
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 YITDGLYATIPGAKPEBEIKRKALEVEYINAKLPGILLEVEYGFYVRGFFVKKRYALI 598
QY 601 DEEGRIITTRGLEVVRDWSIAKETQAKYLEALIKEDSVKAVEIYKDVVEELAKYQVPL 660
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 DEEGKIITTRGLEVVRDWSIAKETQAKYLEALIKGNVEAKIYKEVTEKISKEIIP 658
QY 661 EKLVIHEQITKDSSEYKALGPHVAIAKRLAKIGIKVPRGTIISYIVLRGSGISDRVILL 720
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 659 EKLVIYEQITRPLHEYKALGPHVAIAKRLAARGVYKPRGAVIGIYVLRGDPISKRAIILA 718
QY 721 SEYDPRKKHKYDPDYIENOVLPVAVRLILEAFGRKEDLAKYSSQVGLDAML 772
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 719 EEPDPRKKHKYDAEYIENOVLPVAVRLILEAFGYRKEDLRWOKTKQYGLTAML 770

```

```

CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N
CC PYROPHOSPHATE + DNA(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: AJ250335; CAC12849.1; -.
DR HSSP: p56689; 1TGO.
DR InterPro: IPR002064; DNA_pol_B.
DR Pfam: PF00136; DNA_pol_B; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLBc; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN_1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
SQ SEQUENCE 775 AA; 90417 MW; 633f8392EBE6C73 CRC64;

Query Match          79.8%; Score 3211; DB 1; Length 775;
Best Local Similarity 77.1%; Pred. No. 1,2e-180;
Matches 595; Conservative 86; Mismatches 89; Indels 2; Gaps 1;

QY 1 MFDYDITKDGKPIIRIFKENGEEFKIELDPHPQPIYALALDKSDAIDEIKAKGERHG 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MILDADYITDEGKPIIRIFKENGEEFKVEYDRNFRPIYALLDKDSDQIDEVKKITAEHRG 60
QY 61 KIYRVDAVAVKKKFLGRDVEVKKLFEHPQDYPALRGKIREHPAVIDIYEDIPPAKRY 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 KIYRIVDVEVKKKFLGRPLEVKKLFEHPQDYPALRGKIREHPAVAVDIYEDIPPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEFKGETIMISYADEEARVITKNIDLPY 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEFKGETIMISYADEEGAKVITKKKADLPY 180
QY 181 VDVVSNEREKIKRFVQIVREKDPDLITVNGDNFDPYLKRAEKGVLTLGRDKEHPE 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 VEVVSSEREKIKRFVQIVREKDPDLITVNGDSFDPYLKRAEKGVLTLGRDGS--E 238
QY 241 PKIHRGDSFAVEIKGRHFDLPVVRRTINLPTYLEAVEYALCKTSKGLAEETAI 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 PKMORLGDMTAVEIKGRHFDLVHVRRTINLPTYLEAVEYALFGPKKAVAHETAE 298
QY 301 WETEESMKLAQYSMEDARATYELGKEFFPMEAEALKIGOSVADYSRSSTGNLVEYLL 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 WETGKGLERAKYSMDAKVTYELGREFPMEANQSLVGOPLMDYSRSSTGNLVEYLL 358
QY 361 RAVARNELAPNKPDEBEYRRRLRTTYLGIVKEPERGLMENTIYDFRCLYSIIVTN 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 RKAVERNELAPNKPDEREYERRLRRESYAGGYVEPEKGLMEGLVSDFRSLYPSIIITHN 418
QY 421 VSPDTLEREGCKNVDVAPYIGYKFCDFPFTISIGELITMROEIKKKKATIDPEKK 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 VSPDTLNRBECMEYDVAPEVKKHFCDFPFTISLKLRLDEROEIKRRKAKSDPLEKK 478
QY 481 MLDYRORAVKLANSYGYGYGPKARWYKSCAESYATAMGRHYIEMTIKEIEEKFGEKVL 540
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 MLDYRORAIKILANSYGYGYGAKARWYKSCAESYATAMGREYIEFVRKELEKFGFKVL 538
QY 541 YADTDFYATIPGKPEPTIKKAKKEFLKYINSLPGILLEVEYGFYLRGFFVAKKRAYI 600
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 YITDGLYATIPGAKPEBEIKRKALEVEYINAKLPGILLEVEYGFYVRGFFVKKRYALI 598
QY 601 DEEGRIITTRGLEVVRDWSIAKETQAKYLEALIKEDSVKAVEIYKDVVEELAKYQVPL 660
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 DEEGKIITTRGLEVVRDWSIAKETQAKYLEALIKGNVEAKIYKEVTEKISKEIIP 658
QY 661 EKLVIHEQITKDSSEYKALGPHVAIAKRLAKIGIKVPRGTIISYIVLRGSGISDRVILL 720
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 659 EKLVIYEQITRPLHEYKALGPHVAIAKRLAARGVYKPRGAVIGIYVLRGDPISKRAIILA 718
QY 721 SEYDPRKKHKYDPDYIENOVLPVAVRLILEAFGRKEDLAKYSSQVGLDAML 772
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 719 EEPDPRKKHKYDAEYIENOVLPVAVRLILEAFGYRKEDLRWOKTKQYGLTAML 770

```

RESULT 3

PS Disclosure: Fig 9: 40pp: English.

CC The present invention relates to thermostable mutant B-type DNA
CC polymerases, which have a Y-GG/A amino acid motif between the N-terminal
CC 3'-5' exonuclease domain and the C-terminal polymerase domain, where the
CC lysine of this motif is mutated. The mutant B-type DNA polymerase is
CC useful for synthesizing nucleic acids and for PCR. The present sequence
CC is the protein sequence for a recombinant Thermococcus agarrigans (Tag)
CC DNA polymerase, which was used to illustrate the invention.
XX

Sequence 774 AA:

Query Match 100.0%; Score 4026; DB 22; Length 774;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFDYDITKDGKPIIRIRKENGFEKILDPHOPYITALLKDSADDEIRAKGERHG 60
1 mfdydltkdgpilrirkengfekildphopiyallkdsaddeirakgerhg 60
DB 61 KIVRVDAVKVKKKFGDVEWVKLFEEHPDVPALRGKIREHPAVIDIYEDIPAKRY 120
61 kivrvdavkvkkkfgdvwvklfEEHPDVPALRGKIREHPAVIDIYEDIPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMADIEFTFYHEGDEFGKEIIMISYADEBARVITWKNIDLPY 180
121 lidkglipegdeeklmadiefthyegdefgkeliimsyadebarvitwknidply 180
DB 121 lidkglipegdeeklmadiefthyegdefgkeliimsyadebarvitwknidply 180
QY 181 VDVVSNEREMIKRFQIVREKDPDLITNGDNFDLPYLIRAEKIGVLLGRDKEHPE 240
181 vdvvsneremikrfqivrekdpdlitngdnfdlpyliraeKIGVLLGRDKEHPE 240
DB 181 vdvvsneremikrfqivrekdpdlitngdnfdlpyliraeKIGVLLGRDKEHPE 240
QY 241 PKIHRGDSFAVEIKRIFHFDLPVVRRTINLPTYLEAVYEAVLSEKTSKLGAEETIAI 300
241 pkihrgdsfaveikrIFHFDLPVVRRTINLPTYLEAVYEAVLSEKTSKLGAEETIAI 300
DB 241 pkihrgdsfaveikrIFHFDLPVVRRTINLPTYLEAVYEAVLSEKTSKLGAEETIAI 300
QY 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEKLIGQSVWDVSRSTGNLVEMWYLL 360
301 wetesmkklaqysmedaratyELGKEFFPMEAEKLIGQSVWDVSRSTGNLVEMWYLL 360
DB 301 wetesmkklaqysmedaratyELGKEFFPMEAEKLIGQSVWDVSRSTGNLVEMWYLL 360
QY 361 RYAVENNELAPNKPDEEYRRRLRTYLGYYKPERGMENTYIDFCIYPSIIVTNH 420
361 ryavennelapnkpdeeyrrrlrtylgyykpergmentyidfcIYPSIIVTNH 420
DB 361 ryavennelapnkpdeeyrrrlrtylgyykpergmentyidfcIYPSIIVTNH 420
QY 421 VSPDTEREGCKNYDAPIVGYKCFKDPFGFIPSIIGELITMBOELKKRMKATIDPIERK 480
421 vspdteregcknydapivgykcfkdpfgfipssiIGELITMBOELKKRMKATIDPIERK 480
DB 421 vspdteregcknydapivgykcfkdpfgfipssiIGELITMBOELKKRMKATIDPIERK 480
QY 481 MLDYRGRAVKLHANSYYGYWGYPKARWYSKECAESVTANGRAHYEMTIRKIEEKEFGFVYL 540
481 mldyrgRAVKLHANSYYGYWGYPKARWYSKECAESVTANGRAHYEMTIRKIEEKEFGFVYL 540
DB 481 mldyrgRAVKLHANSYYGYWGYPKARWYSKECAESVTANGRAHYEMTIRKIEEKEFGFVYL 540
QY 541 YATDGFYATIPGEKPEETIKKAKKEFLKYNKSLPGLLELYEGFVLRGFFVAKKRYAVY 600
541 yatdgfyatipgeKPEETIKKAKKEFLKYNKSLPGLLELYEGFVLRGFFVAKKRYAVY 600
DB 541 yatdgfyatipgeKPEETIKKAKKEFLKYNKSLPGLLELYEGFVLRGFFVAKKRYAVY 600
QY 601 DEEGRTITGLEYVRRDWESEIAKETQAKVLEALIKEDSVKAEVIEIKDVEEIAKQVPL 660
601 deegrtitgleyvrrdweSEIAKETQAKVLEALIKEDSVKAEVIEIKDVEEIAKQVPL 660
DB 601 deegrtitgleyvrrdweSEIAKETQAKVLEALIKEDSVKAEVIEIKDVEEIAKQVPL 660
QY 661 EKLVIEHQITKDLSEKKAIGPHVAIAKRLAAKGIKVRPGTIIISYIVLRSGKISDRVILL 720
661 eklviehqitkdlSEKKAIGPHVAIAKRLAAKGIKVRPGTIIISYIVLRSGKISDRVILL 720
DB 661 eklviehqitkdlSEKKAIGPHVAIAKRLAAKGIKVRPGTIIISYIVLRSGKISDRVILL 720
QY 721 SEYDPRKKHYPDYUENOVLPVLELILAEFGYRKEDLVQSSKOGLDAMWKK 774
721 seydpRKKHYPDYUENOVLPVLELILAEFGYRKEDLVQSSKOGLDAMWKK 774
DB 721 seydpRKKHYPDYUENOVLPVLELILAEFGYRKEDLVQSSKOGLDAMWKK 774

RESULT 2
AAW29323
3D AAW29323 standard: Protein: 774 AA.

XX AAW29323;

AC 20-APR-1998 (first entry)

DT DNA polymerase with 3'-5' exonuclease activity.

DE TY Exon; DSM 10597; thermostable; DNA polymerase;
3'-5' exonuclease; amplification.

OS Thermococcus sp.

XX Del19611759-AL.

PD 02-OCT-1997.

PF 25-MAR-1996; 96DE-1011759.

PR 25-MAR-1996; 96DE-1011759.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Antrenkian G, Frey B, Niehaus F;

DR WPI: 1997-48194/45.

DR N-PSDB: AAT86434.

PT Thermostable DNA polymerase from Thermococcus sp. TY - useful for
nucleic acid amplification

PS Claim 1: Pages 9-10; 32pp; German.

CC The present sequence (TY Exon) is a thermococcus sp. TY
(DSM 10597) thermostable DNA polymerase with 3'-5' exonuclease
activity.

CC The enzyme can specifically amplify nucleic acid fragments of up to
5 kb in high yields, has an activity half-life of 20 minutes at 90
degrees C, has an optimum temperature of 70-80 degrees C, has an
optimum pH of 7.5, exhibits optimum activity at a KCl concentration
of 80-100 mM, is magnesium ion-dependent and is inhibited by
manganese ions.

CC Sequence 774 AA:

Query Match 99.1%; Score 3990; DB 18; Length 774;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 769; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MFDYDITKDGKPIIRIRKENGFEKILDPHOPYITALLKDSADDEIRAKGERHG 60
1 mfdydltkdgpilrirkengfekildphopiyallkdsaddeirakgerhg 60
DB 1 mfdydltkdgpilrirkengfekildphopiyallkdsaddeirakgerhg 60
QY 61 KIVRVDAVKVKKKFGDVEWVKLFEEHPDVPALRGKIREHPAVIDIYEDIPAKRY 120
61 kivrvdavkvkkkfgdvwvklfEEHPDVPALRGKIREHPAVIDIYEDIPAKRY 120
DB 61 kivrvdavkvkkkfgdvwvklfEEHPDVPALRGKIREHPAVIDIYEDIPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMADIEFTFYHEGDEFGKEIIMISYADEBARVITWKNIDLPY 180
121 lidkglipegdeeklmadiefthyegdefgkeliimsyadebarvitwknidply 180
DB 121 lidkglipegdeeklmadiefthyegdefgkeliimsyadebarvitwknidply 180
QY 181 VDVVSNEREMIKRFQIVREKDPDLITNGDNFDLPYLIRAEKIGVLLGRDKEHPE 240
181 vdvvsneremikrfqivrekdpdlitngdnfdlpyliraeKIGVLLGRDKEHPE 240
DB 181 vdvvsneremikrfqivrekdpdlitngdnfdlpyliraeKIGVLLGRDKEHPE 240
QY 241 PKIHRGDSFAVEIKRIFHFDLPVVRRTINLPTYLEAVYEAVLSEKTSKLGAEETIAI 300
241 pkihrgdsfaveikrIFHFDLPVVRRTINLPTYLEAVYEAVLSEKTSKLGAEETIAI 300
DB 241 pkihrgdsfaveikrIFHFDLPVVRRTINLPTYLEAVYEAVLSEKTSKLGAEETIAI 300
QY 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEKLIGQSVWDVSRSTGNLVEMWYLL 360
301 wetesmkklaqysmedaratyELGKEFFPMEAEKLIGQSVWDVSRSTGNLVEMWYLL 360
DB 301 wetesmkklaqysmedaratyELGKEFFPMEAEKLIGQSVWDVSRSTGNLVEMWYLL 360

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:17:02 ; Search time 34.46 Seconds

(without alignments)
2494.809 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026
Sequence: 1 MIFDTDTYTKDKGKPIIRIFK.....KEDLYQSSKOYGLDAMLRK 774

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Hit number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /Geneseq_032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4026	100.0	774	22	Recombinant Tag DN
2	3990	99.1	774	18	DNA polymerase wit
3	3446.5	85.6	1829	18	DNA polymerase wit
4	3356	83.4	1702	14	T. litoralis DNA p
5	3232.5	80.3	773	19	Thermococcus gorgo
6	3231.5	80.3	773	19	Thermococcus gorgo
7	3226.5	80.1	774	18	Thermococcus pepto
8	3225.5	80.1	774	18	Heat-resistant DNA
9	3222.5	80.0	774	17	Mature DNA polymer
10	3222.5	80.0	774	18	KOD1 thermostable
11	3221.5	80.0	774	19	Mutant KOD DNA pol

12	3220.5	80.0	774	19	AAW48452	Mutant KOD DNA pol
13	3220.5	80.0	774	19	AAW48458	Mutant KOD DNA pol
14	3219.5	80.0	774	19	AAW48450	Mutant KOD DNA pol
15	3218.5	79.9	774	19	AAW48454	Mutant KOD DNA pol
16	3218.5	79.9	774	19	AAW48455	Mutant KOD DNA pol
17	3218.5	79.9	774	19	AAW48457	Mutant KOD DNA pol
18	3218.5	79.9	774	19	AAW48459	Mutant KOD DNA pol
19	3218.5	79.9	774	19	AAW48460	Mutant KOD DNA pol
20	3218.5	79.9	774	19	AAW48469	Mutant KOD DNA pol
21	3217.5	79.9	774	18	AAW33112	Thermococcus pepto
22	3215.5	79.9	774	19	AAW44731	T. profundus therm
23	3213.5	79.8	774	19	AAW44732	Mutant KOD DNA pol
24	3213.5	79.8	774	19	AAW44732	Mutant KOD DNA pol
25	3211.5	79.6	776	22	AAU03085	Thermococcus DNA p
26	3203.5	79.5	776	22	AAU03075	Thermococcus DNA p
27	3202.5	79.5	776	22	AAU03084	Thermococcus DNA p
28	3202.5	79.5	776	22	AAU03084	Thermococcus DNA p
29	3199.5	79.5	776	22	AAU03086	Thermococcus DNA p
30	3199.5	79.5	776	22	AAU03088	Thermococcus DNA p
31	3199.5	79.5	776	22	AAU03090	Thermococcus DNA p
32	3198.5	79.4	776	22	AAU03081	Thermococcus DNA p
33	3198.5	79.4	776	22	AAU03083	Thermococcus DNA p
34	3198.5	79.4	776	22	AAU03087	Thermococcus DNA p
35	3197.5	79.4	776	22	AAU03076	Thermococcus DNA p
36	3196.5	79.4	776	22	AAU03079	Thermococcus DNA p
37	3196.5	79.4	776	22	AAU03089	Thermococcus DNA p
38	3195.5	79.4	776	22	AAU03078	Thermococcus DNA p
39	3195.5	79.4	776	22	AAU03080	Thermococcus DNA p
40	3195.5	79.4	776	22	AAU03091	Thermococcus DNA p
41	3195.5	79.4	776	22	AAU03092	Thermococcus DNA p
42	3193.5	79.3	774	20	AAW87548	Thermococcus DNA p
43	3192.5	79.3	776	22	AAU03077	Thermococcus DNA p
44	3164.5	78.6	771	22	AAW91133	Thermococcus abyssi
45	3164.5	78.6	774	22	AAW96414	putative P. abyssi

ALIGNMENTS

RESULT 1
ID AAG78941
AC AAG78941; standard: Protein: 774 AA.
XX
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE Recombinant Tag DNA polymerase.
XX
KW Tag polymerase; B-type DNA polymerase.
XX
OS Thermococcus aggregans.
XX
PN EPI132474-A1.
XX
PD 19-SEP-2001.
XX
XX
PF 06-MAR-2001; 2001EP-0104583.
XX
PR 11-MAR-2000; 2000EP-0105155.
XX
PA (HOFF) ROCHE DIAGNOSTICS GMBH.
XX
PI Sobek H, Frey B, Antaniklikian G, Boehlke K, Pilsant FM, Rossi M;
XX
XX WPI: 2001-618367/72.
XX
XX N-PSDB: AAT70053.
XX
XX
XX New thermostable mutant B-type DNA polymerase with a Y-GG/A amino acid
XX motif between the N-terminal 3'-5' exonuclease domain and the
XX C-terminal polymerase domain in the wild type polymerase, useful for
XX synthesizing nucleic acids

XY nucleic acid amplification

DB 721 Vremnikhsynstlrevsseykgevyaltlegnpyytangllthnslpsllvtln 78

```

Db 781 vspdtlercgcknydvaplvgykfckdfpofipsilgelltmrgelkkmkatioplek 840
QY 481 MLDYRQRAVRLHA----- 493
| | | | |
Db 841 mldyrgavrvllansllpnwlpilengevkfvgifidrymeeqdkrvtdntevle 900
QY 494 ----- 493
Db 901 vdnlfafslnkeakkeikvkallrhkygaaeyvelnsgrkhlttrghslftirngk 960
QY 494 ----- 493
Db 961 kelvgeevkvvgdliivpkkvklnekeavlnlpelisklpodedadvmtvtpvkgtrnfk 1020
QY 494 ----- 493
Db 1021 gmlrltlkwlfgeeskrtrtfenrylfhleeljfvklprrgyevtdweglkyrqglyeklvk 1080
QY 494 ----- 493
Db 1081 nlryngnkrreylvrfndikdsvscfprkelewkixlckgfrxkcllkvdedfgkflgyy 1140
QY 494 ----- 493
Db 1141 vsegyagacknktgmsysvklvnenpnvldmknlaekffgvrvygvkncvdlpkkmayl 1200
QY 494 ----- 493
Db 1201 lasklcvtaenkrripallfssaeprwaflrayfvvgdghpskrllstksellangl 1260
QY 494 ----- 493
Db 1261 vflnslgvsiklgfsgsvrvyinedlpflqtrgkntypnllpkveleefgrkfg 1320
QY 494 ----- 493
Db 1321 knltfekfeladsgkldkrkvklldflngdivldrvknvcregyydlsveden 1380
QY 494 ----- 493
Db 1381 flvfglllyahnsygyymgypkarwyskceasvrtawgrnytemtlkeleekfgfkvly 1440
QY 543 D----- 543
Db 1441 dsvtgdtellvkrngtrlefvrpleklfervdyrigekeyclledvealldnrgkllwkv 1500
QY 544 ----- 543
Db 1501 pyvmtrhakkkyrwltnswyldvtedhnlvaedgklearpmelegkallatkddlsq 1560
QY 544 ----- 544
Db 1561 veylkhphaleeisngvyvdlvevegthrfangilvhnrtgfyatlpgepelickkake 1620
QY 566 FLKYINSKLGLELEVEGEFYLRGFFVAKKRYAVIDEGRITRGLSEVVRDMSSEIAKET 625
| | | | |
Db 1621 fllylnsklpjlllelegfyirgffvakkryavidegritrrgllevvrrtdseiaket 1680
QY 626 QAKVLEAILKEDSVEKAVETIKDVEIEIAKYQVPLEKLVINEQITKOLSEYKAIGPHVAI 685
| | | | |
Db 1681 qakvleailkedsevakaveivkdveeialkygpleklyvheqitkdlseykaigphvai 1740
QY 686 AKRLAKGTVRRCTIISYVLRLSGSKISDRVLLSEYOPKHKHYPPDYIENQVPAVL 745
| | | | |
Db 1741 akrlaakgltvrrpctllsyvlrlsgsklsdrvllseydpkknkypdyylenqvlpavl 1800
QY 746 RILEAFGYRKEDLKYOSSKOVGLDAMLKK 774
| | | | |
Db 1801 rllleafigyrkedlkgysskvgldawlk 1829

```

RESULT 4
PAR38095

```

ID AAR38095 standard; Protein; 1702 AA.
XX
AC AAR38095;
XX
DT 21-JAN-1994 (first entry)
XX
DE T. litoralis DNA polymerase.
XX
KW DNA polymerase; intron; translation; bacteriophage NEB #619; E. coli;
KW mutation; cell growth; endonuclease; self splicing; stability;
KW expression; transformation; expression vector.
XX
OS Thermococcus litoralis.
XX
PN EP547920-A.
XX
PD 23-JUN-1993.
XX
PF 18-DEC-1992; 92EP-0311622.
XX
PR 18-DEC-1991; 91US-0811421.
XX
PA (NEMO ) NEW ENGLAND BIOLABS INC.
XX
PI Comb DG, Jack WE, Kucera R, Perlter F;
XX
DR WP1; 1993-198889/25.
XX
DR N-PSDB; AAQ43515.
XX
PT Recombinant thermostable DNA polymerase from archaeobacteria -
PT expressed efficiently by removal of intervening introns
XX
PS Claim 1; Fig 19; 63pp; English.
XX
SS
CC This sequence represents T. litoralis DNA polymerase. The DNA
CC sequence encoding this protein contains introns described in the
CC specification, however this protein sequence appears to be translated
CC through these intronic sequences. Only a portion of the DNA
CC polymerase protein sequence is reproduced in the specification. The
CC T. litoralis DNA polymerase gene sequence may be derived from
CC bacteriophage NEB #619 on an approx. 14kb BamHI restriction fragment.
CC When this gene is transformed in to E. coli and expressed in its
CC entirety, it is found to be unstable. There is a high frequency of
CC mutation in the polymerase gene, cell growth is slow and there is some
CC degree of cell mortality. This instability is due in part to the
CC presence of the two introns. It is thought that the intervening DNA
CC at this sequence encodes a separate protein which may be spliced out
CC at the protein level. The second, 1170bp, intron has been found to
CC code for an endonuclease which self splices out in E. coli. To
CC improve stability of the DNA when it is expressed in E. coli, the
CC specification states that it is desirable to delete the introns
CC before transformation of the expression vector into E. coli.
XX
SQ Sequence 1702 AA;

```

Query Match 83.4%; Score 3356; DB 14; Length 1702;
Best Local Similarity 42.7%; Pred. No. 5e-253;
Matches 726; Conservative 33; Mismatches 15; Indels 928; Gaps 2;

```

QY 1 MFEDDYIRKDKSPITRIIRKNGEKEFTELDHPQPIYALKDSDAIDEIKAIGERHG 60
| | | | |
Db 1 mldtdyitkdgkplritrkngefkeldphptpylalkdsdaaleaklgerhg 60
| | | | |
QY 61 KIVRVDAVAVKVKKLGSDVEVWKLFEHPDPVPLRGKIRHPAVIDIEYDIPFARKY 120
| | | | |
Db 61 kivrldavkvkklfgrevewkllfhpqdvpmrgkltrehpavldyeydipfarky 120
| | | | |
QY 121 LLDKGLIPMEGDEELKNAFDIETFYHSGDERGKEIIMISYADDEBARVITWKINDLPY 180
| | | | |
Db 121 lldkgllpmegeelkllaftietfyhsgdefgkgeilimisysadeearvltwnkndipy 180
| | | | |
QY 181 VDVSNREMERIKRFQIVREKDPDLITVNGNPNFLPLIKRAEKLGVTLGGRKEHPE 240

```

6.

Db	1261	tkqlllpklctyvisnoyphknekqgdlnlaksivkfmkrhfkoeakrrklpefmyelpv	1320
Qy	544	-----	543
Db	1321	tyleaflrglfsadgvtvlrkypelrlnldadflrevrklwlvigslsfaettpnr	1380
Qy	544	-----	543
Db	1381	ynvgstgtyaskhlrlknkwrfaerlgflerqkrllehlksarvkrntldfgfdlvhvk	1440
Qy	544	-----TDFGVARIPGEKPEPTIKKAKELFKAYNS	572
Db	1441	kveipregyvydlveeetrlffamllvhnmdgfyatclpgekpellkkakeflnyins	1500
Qy	573	KLPGLELEEGEFLYLGKGFPAVKRRVAVIDEGSRITTRGILEVVRDMSIAKETQAKYLEA	632
Db	1501	klppllleleyegfylgffvfkrrkyvideegrltrgldevrldwselaketqakylea	1560
Qy	633	ILKEDSYEKRAVELYKQDVVEIAKYQVPLEKLVITHQITRKDLSEYKATGPHVAIKKRLAAK	692
Db	1561	llkxevskavevrvdvveklakryvpleklvlnheqltrdldkdykalgphalakrlaar	1620
Qy	693	GIAKRPETIISYIVLRGSGKISDPRIILSEDPKKHKYDPPDIYIENQVLPVLRITLEAFG	752
Db	1621	gikxkpgtllsylvlkgsgflsdrvlllleqdprrhkhdpdyliengvlpavrlileafg	1680
Qy	753	YRKEDLKYOSSKQGVLDLAWLK	774
Db	1681	yrkedlrygsskqgtjdawlkrl	1702
RESULT 5			
AAW46285	ID	AAW46285 standard; Protein: 773 AA.	
XX	XX	AAW46285;	
DT	28-SEP-1998	(first entry)	
DE	Thermococcus gorgonarius (Tgo) DNA polymerase protein.		
KW	Thermococcus gorgonarius (Tgo) DNA polymerase protein.		
KM	Thermococcus gorgonarius (Tgo) DNA polymerase protein.		
KM	Thermococcus gorgonarius (Tgo) DNA polymerase protein.		
XX	Thermococcus gorgonarius.		
XX	EP834570-A1.		
XX	08-APR-1998.		
XX	03-OCT-1996; 96EP-0115874.		
XX	03-OCT-1996; 96EP-0115874.		
XX	03-OCT-1996; 96EP-0115874.		
XX	(BOEF) BOEHRINGER MANNHEIM GMBH.		
XX	Angerer B, Ankenbauer W, Bonch-osmolovskaya E, Ebenbichler C;		
PI	Laue F, Schmitt-agheguelan G, Svetlichny V;		
XX	WPI: 1998-195468/18.		
DR	N-PSDB; AAV26371.		
XX	New isolated thermostable DNA polymerase - obtained from		
PT	Thermococcus gorgonarius, used for amplifying DNA or for DNA		
XX	cloning, sequencing or labelling		
PS	Disclosure: Page 16-21; 41pp; English.		
CC	This amino acid sequence produces the thermostable DNA polymerase in		
CC	Thermococcus gorgonarius (Tgo). The DNA polymerase enzyme was isolated		
CC	from this species by standard isolation and purification techniques. 7		

CC this activity is not present in Tag polymerase which is normally used in
 CC PCR amplification. Thus this activity enables the fidelity of PCR to be
 CC increased, and therefore would no longer be prone to base incorporation.
 CC The Tgo DNA polymerase exhibits a two fold greater replication fidelity
 CC than any known polymerases. It can also decrease non-specific background
 CC amplification in PCR, due to its 3'-5' activity. The gene from this
 CC species also does not contain intervening sequences, which would have to
 CC be removed to enable expression in *E. coli* to occur. This enzyme has an
 CC approx molecular weight of 92-96 kD, and has been shown to retain 90% of
 CC its activity after incubation, in the presence of a stabilizer for two
 CC hours at 95 deg. C. This Tgo DNA polymerase can be used for e.g.
 CC amplifying DNA or for DNA cloning, sequencing or labelling.

XX Sequence 773 AA:

Query Match 80.3%; Score 3232.5; DB 19; Length 773;
 Best Local Similarity 77.7%; Pred. No. 7.2e-244;
 Matches 601; Conservative 83; Mismatches 86; Indels 3; Gaps 2;

```

1 MFEDDYTKDKGPIIRIKKENGGEFKIELDPHQPPIYALLKDSADIDEIKAKGERHG 60
1 mlltdcyltedgkpyrltkkenggeftldydrnfepylallkdsdpdvkkaetnng 60
61 KIVRVVDKVKKKKFLGRDVEWKLIFEHPOVPALRGKIRHPAVIDIEYDIPFARKY 120
61 tlvrvvraekvkfkflgrplevwklyftbpdvpalrdkikhpnvavdiyeydipfarky 120
121 LIDKGLIPMEGDEELKLMFDIETFYHSGDEFGKEIIMISYADEBARVITWKNDLPY 180
121 lldkglipmegdeeklmfadietfyehegeefaeagplmisyadeegarvltwnkndipy 180
181 VDVSNEREMIKRFVOIYREKDPDLITNGDNFDPYLIRAKELGVTLLGRDKEHPE 240
181 vdvsnereimikrfvoiyrekdpsdlitngdnfpylirakekgvtllgrdkehpe 240
241 PIHMGDSFAVEIKRIFEDLFPVVRRTINLPTYLLEAVEIRGAKTSKIGAEIRAI 300
241 pihmgdsfaveikrifeidlfpvvrirtinlptylleaveirgaksksigaeirai 300
239 pldgtmgdfavevgrifldlyprlritinlpylleavayallgqbrkayaeelaqa 298
301 WETESMKRLAOYSMEDARATYELGKEFPMEAEIAGISQWMDSRSSSTGNLYEWYLL 360
301 wetesmkrlaoysmedaratyelgkefpmeeaelagisqswmdsrssstgnlyewyll 360
299 wetgseglervarysmedavelyelgkeffmeaqslrlyvgsltwdsrstgnlyewfll 358
361 RYATERNELAPKPKDEEYERRRLRTTYLGGYKKEPREGIMENITLIDFCLYPSIIVTN 420
361 ryaternelapkpdeeyerrrlrttylgyykekepergimennitlidfclypsiiivtn 420
359 rkayernelapkpderelair-resyaggykpeperglwenlyvldfrslypsliithn 417
421 VSPDTEREGCKNNDYAPIVGKFCDFPGFIPSIIGELITRHOELTKKMKATIDPIEEK 480
421 vspdteregcknnydypivgkfcdfpgfipsiigelitrhoeltkkmmkatidpieek 480
418 vspdclnregeceydvapgvgnhctcdipglfipslilgdlleerqvkmmkatalidpiekk 477
481 MLDYRQAVAKLHANSYYGYGPKARWYSEKCAESVTANGRYIEMTIEIEKEGFKVL 540
481 mldyrqavaklhansyygygpkarwysckcaesvtangryiemtieiekegfkvl 540
478 lldygralkilansfygygyltkarwykcaesvtangryiemtieiekegfkvl 537
541 YADTDFATITGEKPEITKKAKERFLKYNKSLPGLLELEHGEFLTKRFFPAKKRYAVI 600
541 yadtdfatitgeketitkkakerflkynkslpgllelehegefltkrffpaakryavi 600
538 yadltgdfatitpdaetkakkakelfdyinaklpjllleleyegfkyrgfivclkkkyavi 597
601 DEEGRTTGTGLEVRNDSMEIKETQAKVLEALILKEDSVKEVEIKVDVIEAAKQVQL 660
601 deegrttgtglevnrndsmeketqakvlealilkedsvkeveikvdvieaakqvql 660
598 deedkrltctgtlelvtrrdweikeqcarylealikhgdeeevrlkveteklskyevpp 657
661 EKLVIHQETKDLSEYKAIGPVAIALAKRLAAGIKVRPCTIISYVLNCSGISDPVILL 720
661 eklviheqetkdlseykaigpvaialakrlaagikvrpctiisylvlnscgisdpvill 720
658 eklviyegrltrldkqkagpnhavakariaegiklrptvlsyvtlksqslgdratrf 717
721 SEYDPKHKHYDDPYITENOVLPVAILERLFAFGYRKEDLKYQSSKQVGLDAMITK 773
721 seydpkhhkyddpyitenovlpvailerlfafgyrkedlkyqsskqvgladamitk 773
718 defdpakhhkydaeyylengvlpaverllrafgyrkedlkyqtrgyglgawik 770

```

RESULT 6
 AAM58818
 ID AAM58818 standard; Protein; 773 AA.
 AC AAM58818;
 DT 17-AUG-1998 (first entry)
 DE Thermococcus gorgonarius thermostable DNA polymerase.
 XX DNA polymerase; thermostable enzyme; DNA amplification; PCR;
 KW cloning; sequencing; DNA labelling.
 OS Thermococcus gorgonarius.
 PN EP834571-A1.
 XX 08-APR-1998.
 PD 16-JAN-1997; 97EP-0100584.
 PF 03-OCT-1996; 96EP-0115874.
 PR (BOEHR) BOEHRINGER MANNHEIM GMBH.
 PA Angerer B, Ankenbauer W, Bonch-Osmolovskaya E, Ebenbichle RC;
 PI Laue F, Schmitz-Agnehan G, Svetlichny V;
 DR WPI: 1998-195469/18.
 XX N-PSDB: AAV11396.
 PT New isolated thermostable DNA polymerase - obtained from
 PT Thermococcus gorgonarius, used for amplifying DNA or for DNA
 PT cloning, sequencing or labelling.
 PS Claim 1; Fig 3; 41pp: English.

CC This polypeptide comprises a novel thermostable DNA polymerase
 CC (TDP) of Thermococcus gorgonarius, a thermophilic archaeobacterium
 CC isolated from a thermal vent. The TDP catalyses the directed
 CC polymerisation of DNA and retains about 90% of its activity after
 CC incubation for 2 hr at about 95 deg C in the presence of a
 CC stabilizer. It exhibits more than 2-fold greater replication
 CC fidelity than DNA polymerase from *Pyrococcus furiosus*, and also has
 CC 3'-5' exonuclease (proofreading) activity. This exonuclease
 CC activity can decrease non-specific background amplification in PCR
 CC by degrading delayed ends of primers bound to unspecific
 CC sequences. The TDP is suitable for use in DNA amplification,
 CC cloning, sequencing and labelling. The native enzyme can be used,
 CC or a recombinant enzyme that has been expressed in microbial
 CC host cells transformed with a vector carrying an isolated DNA
 CC sequence (see AAV11396) coding for TDP.

XX Sequence 773 AA:

Query Match 80.3%; Score 3231.5; DB 19; Length 773;
 Best Local Similarity 77.7%; Pred. No. 8.6e-244;
 Matches 601; Conservative 83; Mismatches 86; Indels 3; Gaps 2;

```

1 MFEDDYTKDKGPIIRIKKENGGEFKIELDPHQPPIYALLKDSADIDEIKAKGERHG 60
1 mlltdcyltedgkpyrltkkenggeftldydrnfepylallkdsdpdvkkaetnng 60
61 KIVRVVDKVKKKKFLGRDVEWKLIFEHPOVPALRGKIRHPAVIDIEYDIPFARKY 120
61 tlvrvvraekvkfkflgrplevwklyftbpdvpalrdkikhpnvavdiyeydipfarky 120
121 LIDKGLIPMEGDEELKLMFDIETFYHSGDEFGKEIIMISYADEBARVITWKNDLPY 180
121 lldkglipmegdeeklmfadietfyehegeefaeagplmisyadeegarvltwnkndipy 180
181 VDVSNEREMIKRFVOIYREKDPDLITNGDNFDPYLIRAKELGVTLLGRDKEHPE 240

```

```

Db      181 vdvastekemikrfikvvkekdpvlllyngdnfdafaylkrseklygkfllygr--egse 238
Oy      241 PRHMGSPFAVEIKRGRIHFDLFPVVRRTINLPTTLEAVYAVANGKRSKIGAEINAI 300
Db      239 plqigmqrfaevvgrihfdlyprlrlinplytleavyaallgqkkekyaeeelaqa 298
Oy      301 WETESMKLAQYSMEDARATYELGKEFFPMEAEALKIGOSVMDVSSSTGNLEWYLL 360
Db      299 wetseglervarvamedekvlyelgkeffpmeaqlsrllvgqslwvsvssgnlvefll 358
Oy      361 RYAVRNELAPNKPDEEYRRRLRTTYLGIVYKEPERGMENTYLDPRCLYPSIIIVTNH 420
Db      359 rkayernelapnkpdekelarr-rsyyagvykpeperglwenlvyldfrslypsllthn 417
Oy      421 VSPDLREERGCKNVDVAVIVGKFKDPGFPISILGELITMRORIKKMKATDPIREK 480
Db      418 vspdlmregceeyvavpvgnhkcdkdfpdlpsllgdlleergkvrkkmkaldpblek 477
Oy      481 MLDYRORAVKLANSYGYMGYPKARWYSKECAESVTAMGRHYIMDTIKEIEKGFYVL 540
Db      478 lldygrariklansfygygrytkarwykceaeesvtmgreyietllreleekfgfvl 537
Oy      541 YADTGFYATIPGKEPETIKKAKFEFLKYINSKLPGLLELEYEGFYLGPFVAKKRYAVI 600
Db      538 yadtgffatipgdaaevtvkkakeflidylnaklpgllleleyegfykrgffvtkkkyavi 597
Oy      601 DEEGRTTRGLEVVRDMSEIAKETQAKVLEALIKEDSVKRAVEIVKRYVEIAKYQVPL 660
Db      598 deegrttrglelvrrdmseiakeqarvleallkbgdveevrlykveteklskyevpp 657
Oy      661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVKRGTTISYIVLGSCKISDRVILL 720
Db      658 eklviheqitrdlsdykatgphvavakrlaargiklrpgvtvlsyvlkysgrlgralpf 717
Oy      721 SEYDPKHKHYDPDYIENQVLPVAVRLILEAFGRKEDLKYOSSKQVGDAMWLK 773
Db      718 defdpknhkyaeyylengvlpaverllraifgyrkedlrygktrqvglsavlk 770

RESULT 7
AAW33111
ID      AAW33111 standard; Protein: 774 AA.
XX
AC      AAW33111:
XX
DT      28-JAN-1998 (first entry)
XX
Dy      Thermococcus peptonophilus OG-1 thermostable DNA polymerase.
XX
Kw      JCM 9653; thermostable; DNA polymerase;
XX      nucleic acid sequence amplification; PCR.
XX
OS      Thermococcus peptonophilus.
XX
PN      JF09252776-A.
XX
PD      30-SEP-1997.
XX
PE      19-MAR-1996; 96Jp-0063112.
XX
PR      19-MAR-1996; 96Jp-0063112.
XX
PA      (TOYM ) TOYOBO KK.
XX
DR      WPI: 1997-530149/49.
XX      N-PSDB; AAT88373.
XX
PT      Thermococcus peptonophilus thermostable DNA polymerase - useful for
XX      nucleic acid sequence amplification, e.g. polymerase chain reaction
XX
PS      Claim 9; Pages 9-12; 27pp; Japanese.
XX

```

```

CC      The present sequence is a Thermococcus peptonophilus OG-1
CC      (JCM 9653) derived thermostable DNA polymerase, which can be used
CC      for nucleic acid sequence amplification, e.g. PCR.
XX
SQ      Sequence 774 AA:

Query Match      80.1%; Score 3226.5; DB 18; Length 774;
Best Local Similarity 77.7%; Pred. NO. 2.1e-243;
Matches 601; Conservative 82; Mismatches 87; Indels 3; Gaps 2;

Oy      1 MIFDDYITKDKPIIRIFKKNGEKFIEIDPHFOPYIYALAKDSADIEKAIGERHG 60
Db      1 mlldtdytlengkprlirifkkngefkleydtcfepyllyallkdsaleevkkltaehg 60
Oy      61 KIVRVDAVKKKKKFLGSDVDEWKLIFFEHPODVPALRGKIRHPAVIDIYEYDIPAKRY 120
Db      61 tvltvkraekvekkflgprlewkllyfthpgdvpaalrdkirehpavidlyeydipfakry 120
Oy      121 LIDKGLIMEGDEELKMAFDIETFYHGDDEFGKEITIMISVADDEEARVITMKINDLPY 180
Db      121 lldkglimegedelkmafdietlyhegeetaegppllmisvadeegarvltwnldppy 180
Oy      181 VDVVSNREMIKRFVOIVREKDDVLLITYNGDNFDLPYLIRKRAEKLGYTLLEGDRKEHP 240
Db      181 vdvvsteremikrfvriivkekdpvlllyngdnfdafaylkrckeklglnfalgrds--e 238
Oy      241 PRHMGSPFAVEIKRGRIHFDLFPVVRRTINLPTTLEAVYAVANGKRSKIGAEINAI 300
Db      239 plqigmqrfaevvgrihfdlyprlrlinplytleavyaallgqkkekyaeeelaqa 298
Oy      301 WETESMKLAQYSMEDARATYELGKEFFPMEAEALKIGOSVMDVSSSTGNLEWYLL 360
Db      299 wetseglervarvamedekvlyelgkeffpmeaqlsrllvgqslwvsvssgnlvefll 358
Oy      361 RYAVRNELAPNKPDEEYRRRLRTTYLGIVYKEPERGMENTYLDPRCLYPSIIIVTNH 420
Db      359 rkayernelapnkpdekelarr-rsyyagvykpeperglwenlvyldfrslypsllthn 417
Oy      421 VSPDLREERGCKNVDVAVIVGKFKDPGFPISILGELITMRORIKKMKATDPIREK 480
Db      418 vspdlmregceeyvavpvgnhkcdkdfpdlpsllgdlleergkvrkkmkaldpblek 477
Oy      481 MLDYRORAVKLANSYGYMGYPKARWYSKECAESVTAMGRHYIMDTIKEIEKGFYVL 540
Db      478 lldygrariklansfygygrytkarwykceaeesvtmgreyietllreleekfgfvl 537
Oy      541 YADTGFYATIPGKEPETIKKAKFEFLKYINSKLPGLLELEYEGFYLGPFVAKKRYAVI 600
Db      538 yadtgffatipgdaaevtvkkakeflidylnaklpgallleleyegfykrgffvtkkkyavi 597
Oy      601 DEEGRTTRGLEVVRDMSEIAKETQAKVLEALIKEDSVKRAVEIVKRYVEIAKYQVPL 660
Db      598 deegrttrglelvrrdmseiakeqarvleallkbgdveevrlykveteklskyevpp 657
Oy      661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVKRGTTISYIVLGSCKISDRVILL 720
Db      658 eklviheqitrdlsdykatgphvavakrlaargiklrpgvtvlsyvlkysgrlgralpf 717
Oy      721 SEYDPKHKHYDPDYIENQVLPVAVRLILEAFGRKEDLKYOSSKQVGDAMWLK 773
Db      718 defdpknhkyaeyylengvlpaverllraifgyrkedlrygktrqvglsavlk 770

RESULT 8
AAW41313
ID      AAW41313 standard; Protein: 774 AA.
XX
AC      AAW41313:
XX
DT      20-MAY-1998 (first entry)
XX
Dy      Heat-resistant DNA polymerase.
XX

```

```

XX Heat-resistant DNA polymerase; 3'-5' exonuclease activity;
KW nucleic acid amplification.
XX
XX Pyrococcus sp.
OS
FH Key Location/Qualifiers
FT Misc-difference 140..144
FT /note="motif mutated to give polymerase of the
FT Invention"
XX
XX JPI0042871-A.
XX
XX 17-FEB-1998.
XX
XX 29-JUL-1996; 96JP-0198910.
XX
XX 29-JUL-1996; 96JP-0198910.
XX
XX (TOYO ) TOYOBO KK.
XX
XX WPI: 1998-186867/17.
XX N-PSDB: AAV14524.
XX
XX New heat-resistant DNA polymerase - useful for, e.g. increasing
XX efficiency of nucleic acid amplification
XX
XX Claim 3: Page 17-19; 22pp; Japanese.
XX
XX This sequence represents the wild-type heat-resistant
XX DNA polymerase of the invention (HRDP), which has a 3'-5' exonuclease
XX activity. In the HRDP of the invention, the motif shown in AAV41315 is
XX mutated. The enzyme can increase the amplification efficiency, of
XX nucleic acid amplification methods.
XX
XX Sequence 774 AA:
SO

```

Query Match 80.1%; Score 3225.5; DB 19; Length 774;
 Best Local Similarity 77.5%; Pred. No. 2,5e-243;
 Matches 599; Conservative 82; Mismatches 89; Indels 3; Gaps 2;

```

QY 1 MFEDPTIRKDKPRIRIRKKEGKEFKELDPHFOPRYVALKKDSADIEIKAKGERRG 60
DB 1 mldtlytedgkprvrlrtkkngefkleydrfeyfyalldksaleevkklaetrg 60
QY 61 KIYRVVDAVKKKKFLGRDVEWKLIFEHODYPALRGKIREHPAVIDIYEDIPFAKRY 120
DB 61 tvvtvkrvekvqkklfgprvewklyftnpgdvalrtdkirehgvldlyeydlyfakry 120
QY 121 LIDKGLIPMEGBELKMAFDIEFTYHEDGDFEGKEIIMISVADDEARVYTKNIDLPY 180
DB 121 lldkglypmegdeelmldfietlyhegeefaeqpllmisyadeagavltlknvdlpy 180
QY 181 VDVVSEMERIKRFOIVREKDPDLITNGDNFDLPYLIRAKRKGVTLLGRDKEHP 240
DB 181 vdvvsereimkrflrvvnekdpdvlitngdnfdlaylkrcekignfaylgrdgs--e 238
QY 241 PKIHRGDSFAVEIKIRIFDLFPVVRTINLETTYLLEAVENAVLTKSLGAEBIAAI 300
DB 241 pkirngdrfavevkrhfdlyprvrrtlnlptyleavenvfqpkekyaeeltpa 238
QY 301 WETESMKLAOYSMEDARATYELGKEFFPMELAKLIGOSWDVSRSTGMLVWYLL 360
DB 301 wetesmklaoysemadaratylelgeffpmeaelakligoswdvsrstgmlvwyl 358
QY 361 RAVYENELAPNKPDEEVRRRRTYIGGYVEPERGMLMENTYIDPRLYSSIVTNH 420
DB 361 ravyenelapnkpdeevrrrrtylgygyvepergmlmentyidprrlyssivtnh 417
QY 421 VSPDTEREGCKNVDAPYVYKFCDFPGFIPISLIGELITMRQETIKKKAKATIDIEKK 480
DB 421 vspdteregcknvdapvyvyrkfcdfpgfipisligelitmrqetikkakatidiekk 477
QY 418 vepdclnregckeydvapvyghrfckdfpgfipisligdlleerqklkkakatidpker 477

```

```

QY 481 MDYRORAVKLANSYGYGYPKARWSECAESVTANGRHYIENTINEIEKFGKVL 540
DB 478 lldygralkllansygygyararwykeceesvltangreytlmtlkeiekyfklv 537
QY 541 YADTGSFYATTIGKEKPTTKKAKKEFLKYINSKLPLELEEGFLRQFPFAKKRYANI 600
DB 538 ysdtdgltfalfpdaetevkkameflnylnaklpaleeyegfkyrgfivlkkkyavl 597
QY 601 DEGRITTRGLEVRDRDWEIARETOAKVLEAIKDSVEKAVEIKDVVEETAKYQVPL 660
DB 598 deegklttrglvdrdweiseiketgarvleallkdgdevkavrlvkevtelksyevp 657
QY 661 EKLVIHEQITKDLSEYKAGPVAIAKRLANGIKYRPGTIIISYLVKSGKISDRVILL 720
DB 658 eklvheqitrdkldykacqphvavakrlaargvklrpgtvisylvksgygrdralpf 717
QY 721 SEYDPKKHKYDDPYRYENOVAPVLRILAPFYGRKRDLDKQSSKQVGLAWLK 773
DB 718 detdpckhkyyddpyryenovlpraverilrafygrkedlryqktrqyglawlk 770

```

RESULT 9
 AAR93154
 ID AAR93154 standard; Protein; 774 AA.
 XX
 AC AAR93154;
 XX
 DT 20-JUN-1996 (first entry)
 XX
 DE Mature DNA polymerase from superthermophilis archaeobacterium KOD1.
 XX
 XX DNA polymerase; superthermophilic archaeobacterium; amplification; primer;
 KW PCR; Pyrococcus furiosus; probe; E.coli; expression vector;
 XX thermal stability.
 XX
 OS Synthetic.
 XX
 XX JP07298879-A.
 XX
 XX 14-NOV-1995.
 PD
 XX
 PF- 09-MAY-1994; 94JP-0095109.
 XX
 XX 09-MAY-1994; 94JP-0095109.
 PR
 XX
 PA (TOYO) TOYOBO KK.
 XX
 PA WPI: 1996-026421/03.
 DR
 XX
 XX DNA polymerase from a super-thermophilic archaeobacterium - is
 PT suitable for polymerase chain reaction
 PS
 PS Claim 4: Page 12-14; 20pp; Japanese.
 XX
 CC This is the amino acid sequence of the mature DNA polymerase from the
 CC superthermophilic archaeobacterium KOD1. The protein was constructed
 CC from the complete gene (AAT10258) by removing the 2 intervening
 CC sequences (nucleotides 1374-2453 and 2708-4316) by fusion PCR using the
 CC primers AAT10262-7. The resultant fragment was cloned into the
 CC expressionvector pET-8c to produce pET-pol for expression and
 CC purification of the protein in E.coli.
 CC The complete gene was isolated by amplifying a fragment (AAT10261) using
 CC primers AAT10259-60 which were based on conserved amino acid sequence
 CC from the Pyrococcus furiosus DNA polymerase. The amplified fragment was
 CC used as a probe to isolate the full length gene from a Southern
 CC hybridisation on chromosomal DNA. The gene was cloned into plasmid pBS
 CC and used to transform E.coli JM109 from which a plasmid contg. the full
 CC length gene was isolated. Cells contg. the recombinant expression
 CC plasmid were used to express and purify the KOD1 DNA polymerase which is
 CC suitable for PCR due to its high thermal stability.
 CC
 SO Sequence 774 AA.

Query Match 80.0%; Score 3222.5; DB 17; Length 774;
 Best Local Similarity 77.4%; Pred. No. 4.4e-243;
 Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

```

QY 1 MIFPDYITKDGKPIIRIFKKNKNGEFKTELDPHFOPIYALIKDSDAIDEIKAIKGERHG 60
DB 1 mldcdyitdegkpyirifkknkngfkyeyrtfeyfayllkdsaleevkkltaerhg 60
QY 61 KIVRVDAVKKKKKFLGDNVEWKLIFEHPODVPALRKIRHRAVIDIYEDIPFARKY 120
DB 61 lvtvkrvekyqkklfgrvewwkllyfchpqdpalrcklrehgavldiyedipfarky 120
QY 121 LIDKGLIPMEGDEELKLAFDIETFYHGEDEFGKCEIIMISYADEEARVITWKNDIPY 180
DB 121 lldkglipmgedeelkmlafidqcltlyhegeefaeegpllmisyadeegarvltwkndipy 180
QY 181 VDVSNEREMIKRVOYIAREKDPDLITYNGDNFDLPYLIRAKELGVTLLGRDKERPE 240
DB 181 vdvsteremikrfirvkekdpdvlityngdnfdlaylkrckelglnfalgrdgs--e 238
QY 241 PKIRMGDSFAVEIKGRHFDLPFVVRRTINLPYTLAEVAVLGKTKSKIGAEIATAI 300
DB 241 pkirmgdsfaveikgrhfdlpfvrvtinlpytleaveavlgqpkkevyaeeelpa 298
QY 301 WETESMKLAQYSMEDARATYELGKEFFPMEAEELAKLIGOSVMDVSSTGNLWEMYL 360
DB 301 wetenleravarymedekvyeigkelflmeaqsrllgslwvdsrsgnlwefll 358
QY 361 RVATERNELAPNKPDDEEYRRRLRTTYIGGYKPEREKLMENTIYLDRCCLPSIIVTHN 420
DB 361 rkayernelapnkpdekeltarr-rqsyeggyvkeperglwenvlyldfrslypsllthn 417
QY 421 VSPDLEREGCKNTDVAIVGKFKCKDPGFPISILGELIMROIKKKMATIDPIRKK 480
DB 418 vspdlmregckvavpqvghrtckdfpflpsllgdlleerqkikkmatldpierrk 477
QY 481 MLDVROARAVKLHANSYGYMGYPRKARWYSKCAESVTAMGHYIMTITKEIEKGFVYL 540
DB 478 lldvrgaravklhansyygyrararawyckeaeesvtamgryeilmtekeleekygfvi 537
QY 541 YADPDGEYATTPGEKPEITIKKAKEELKYINSKLPGLELEYEGFYLRGFPVAKKRVAVI 600
DB 538 ysdctgffactlpgdaaevtvkkameflnylnaklpgaleleyegfykrqgffvckkkyavi 597
QY 601 DEBERITTRGLEVVRKMSSEIAKETQAKVLEALILKEDSEKAVELVKNVVEIAKYQVPL 660
DB 598 deegkilttrgleivrrdseialeketqarvleallkdgdvkavrlvkevteliskyevp 657
QY 661 EKLVIHQITKDLSEYKALIGPHVAIAKRLAKGIVRGRTIISYVLGSGKISDRVILL 720
DB 658 eklviheqitrdlkdykactlpgbhavakrlaargvklrpgvtlisyvllgsgrtigdrtafpf 717
QY 721 SEYDPKKHAYDPDYIENQVLPVLRILAEFYRKREDLKYGSSQKQGLDAMLK 773
DB 718 defdpkthkyaeyyienuvlpaverllraefgylkedlrygktrqvgysawlk 770

```

RESULT 10

AAW20048 ID AAW20048 standard; Protein: 774 AA.

AAW20048; AC

05-SEP-1997 (first entry)

KOD1 thermostable DNA polymerase having 3' to 5' exonuclease activity.

KOD1: archaeon: archaebacteria; thermostable DNA polymerase;

KW heat resistant; intervening sequence; amplification; PCR;
 KW polymerase chain reaction; 3' to 5' exonuclease activity;
 KW hyperthermophilic.

```

XX XX Hyperthermophilic archaeon strain KOD1.
OS XX
XX XX EP745675-A2.
XX XX
XX XX 04-DEC-1996.
XX XX
XX XX 30-MAY-1996; 96EP-0108613.
XX XX
XX XX 31-MAY-1995; 95JP-0134096.
XX XX
XX XX (TOYM ) TOYO BOSEKI KK.
XX XX (TOYM ) TOYOBO KK.
XX XX
XX XX Arakawa T, Imanaka T, Inoue H, Kawakami B, Kawamura Y;
XX XX Kitabayashi M, Morikawa M, Takagi M;
XX XX WPI: 1997-013698/02.
XX XX N-PSDB: AAT71296.
XX XX
XX XX New isolated thermostable KOD polymerase - from hyperthermophilic
XX XX archaeon strain KOD1, used for amplifying target nucleic acids
XX XX
XX XX Claim 3; Page 12-15; 44pp; English.
XX XX
XX XX AAW20049 is a thermostable DNA polymerase (TDP) derived from the
XX XX hyperthermophilic archaeon strain KOD1. The DNA sequence encoding this
XX XX enzyme contains intervening sequences, which are removed to leave a
XX XX sequence encoding the polymerase. This sequence is a truncated version
XX XX of the entire product of the gene (see AAW20048). The TDP has a DNA
XX XX extension rate of at least 30 bases per second and a 3'-5' exonuclease
XX XX activity. The TDP is used for amplifying target nucleic acid with a
XX XX short reaction time and high fidelity.
XX XX
XX XX Sequence 774 AA:

```

Query Match 80.0%; Score 3222.5; DB 18; Length 774;
 Best Local Similarity 77.4%; Pred. No. 4.4e-243;
 Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

```

QY 1 MIFPDYITKDGKPIIRIFKKNKNGEFKTELDPHFOPIYALIKDSDAIDEIKAIKGERHG 60
DB 1 mldcdyitdegkpyirifkknkngfkyeyrtfeyfayllkdsaleevkkltaerhg 60
QY 61 KIVRVDAVKKKKKFLGDNVEWKLIFEHPODVPALRKIRHRAVIDIYEDIPFARKY 120
DB 61 lvtvkrvekyqkklfgrvewwkllyfchpqdpalrcklrehgavldiyedipfarky 120
QY 121 LIDKGLIPMEGDEELKLAFDIETFYHGEDEFGKCEIIMISYADEEARVITWKNDIPY 180
DB 121 lldkglipmgedeelkmlafidqcltlyhegeefaeegpllmisyadeegarvltwkndipy 180
QY 181 VDVSNEREMIKRVOYIAREKDPDLITYNGDNFDLPYLIRAKELGVTLLGRDKERPE 240
DB 181 vdvsteremikrfirvkekdpdvlityngdnfdlaylkrckelglnfalgrdgs--e 238
QY 241 PKIRMGDSFAVEIKGRHFDLPFVVRRTINLPYTLAEVAVLGKTKSKIGAEIATAI 300
DB 239 pkirmgdsfaveikgrhfdlpfvrvtinlpytleaveavlgqpkkevyaeeelpa 298
QY 301 WETESMKLAQYSMEDARATYELGKEFFPMEAEELAKLIGOSVMDVSSTGNLWEMYL 360
DB 299 wetenleravarymedekvyeigkelflmeaqsrllgslwvdsrsgnlwefll 358
QY 361 RVATERNELAPNKPDDEEYRRRLRTTYIGGYKPEREKLMENTIYLDRCCLPSIIVTHN 420
DB 359 rkayernelapnkpdekeltarr-rqsyeggyvkeperglwenvlyldfrslypsllthn 417
QY 421 VSPDLEREGCKNTDVAIVGKFKCKDPGFPISILGELIMROIKKKMATIDPIRKK 480
DB 418 vspdlmregckvavpqvghrtckdfpflpsllgdlleerqkikkmatldpierrk 477

```


XX	AA04852	standard; protein; 774 AA.
XX	AA04852:	
XX	15-JUL-1998	(first entry)
XX	Mutant KOD DNA polymerase (210ND).	
XX	KOD DNA polymerase; 3'-5' exonuclease activity; amplification.	
XX	Pyrococcus sp.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
XX	Misc-difference 210	/note- "Changed from Asn in wild-type to Asp in mutant"
XX	Domain	137..146
XX	Domain	/note- "Exo 1 region"
XX	Domain	206..222
XX	Domain	/note- "Exo 2 region"
XX	Domain	308..318
XX	Domain	/note- "Exo 3 region"
XX	EP822256-A2.	
XX	04-FEB-1998.	
XX	24-JUL-1997;	97EP-0112760.
XX	30-JUL-1996;	96JP-0200446.
XX	29-JUL-1996;	96JP-0198911.
XX	(TOYM) TOYO BOSEKI KK.	
XX	Imanaka T, Kamimura H, Kawakami B, Kawamura Y, Kitabayashi M;	
XX	Komatsubara H, Takagi M;	
XX	WPI; 1998-102909/10.	
XX	Modified thermostable DNA polymerase - with reduced 3'-5' exonuclease activity	
XX	Claim 13; Page -: 54pp; English.	
XX	The invention claims a mutated thermostable KOD DNA polymerase of the present sequence. It has 0.1% of the 3'-5' exonuclease activity when compared with the wild-type KOD DNA polymerase. It also has a DNA extension rate of at least 30 bases/sec, while retaining at least 60% of its activity after heating at 95 deg. C for 6 hours. The invention also claims other mutated versions of the polymerase (AA04849-W48461). The invention shows that efficiency in DNA amplification can be achieved by using a mixture of two different DNA polymerases which are almost identical to each other with respect to thermostability and DNA extension rate but are different in their 3'-5' exonuclease activity. The mixture may contain the present polymerase with a second one which may be the wild-type KOD polymerase or a mutant polymerase (AA04845-W48461) having 6-100% of the 3'-5' exonuclease activity of a native KOD polymerase. N.B. The present sequence is not given in the specification, but is made up using the given wild-type KOD polymerase sequence and the information provided in the claims section.	
XX	Sequence 774 AA;	

Query Match	80.0%;	Score 3220.5;	DB 19;	Length 774;
Best Local Similarity	77.4%;	Pred. No. 6.3e-243;		
Matches 598;	Conservative 83;	Mismatches 89;	Indels 3;	Gaps 2

```

Oy 1 MIEFDYITKQKPIRIFFKKEEKFIEDPHFOPYIYALLKDDSAIDEIKAIKGERHG 600
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 1 mldcdyitegkpvirifkknegfkiedrtfepfyallkddsaaleavkkaierhg 600

```

QY	61	KIVRVLVDAUVKXKKKFLGSDVWMLLEGNRODVALLGKIRREHNAVUIDEYODRPRKKY	120
Dd	61	twcvtkivexqvkqfllgrvwevwyklytltbpdvralttoklrebnavldIeydIpraktu	120
QY	121	LIDKGLJMEGDEELKLMAFDIEFFYHEGDEFGGELIMISTYADEEARVITMKNDLRY	180
Dd	121	lIdKglJmEGdeElKlMafdieffYheGdeFGGElImIStYadEEARvITmKNDlRy	180
QY	181	VDVYNEEEMKIKRFQOYIKREKDRPVLYITTYNDNDNDRLYLKIRAKLGVTLILLRDKKNRE	240
Dd	181	vDvYnEEeMKIKrFQOyIKrEKDrPVLYITTYNDNDNDRLyLKIRAKLGVTLILLRDKKNRE	240
QY	241	PKIHNRGSPFAVEIKGRITHPDLRFVVRRTIMLPIYTLEAVYEAVALGKTKSGAEEETAAI	300
Dd	239	pKIHnrgSPfAVEIKGRITHpDLRFVVRRTIMLPIyTLEAVYEAVALGKTKSGAEEETAAI	298
QY	301	WETESMAMKAOYSMEDARATYELGKEFFRHEBALALIGOSWADVSRSTGMLYEVYLL	360
Dd	299	weTegSMAmKAOySMEDARATyELGKEFFRHEBALALIGOSwADVSRSTGMLYEVYLL	358
QY	361	RYAVERNELAENKRDDEEYERRLRTTYTLYGCVKPREGLMENTYLDPRCLYPSIITVNH	420
Dd	359	rYAvERnELaENKRDDEEYERRLRTTYTLYGCVKPREGLMENTYLDPRCLYPSIITVNH	417
QY	421	VSPDTLEEGCKKNIDVAPITVIGYKCKDPPGTSIBSELITMRQELKXKKKATIDPREKK	480
Dd	418	vSPdTLeeGCKKNIDVAPITVIGYKCKDPPGTSIBSELITMRQELKXKKKATIDPREKK	477
QY	481	MLDYKORAVKSHANSYUYMYGMPKPARWYKSCASVAMGNYHEMJKETEKEFGFKVLL	540
Dd	478	lIdYgRvAikShANSyUYyUZataryKcKeeASvAMgNyeJtmJketeEygfkVl	537
QY	541	YADDDGFATIPGEKPEITIKKAKAEFLYINSKLRPLLELEEFYLRGFFVAKKRRAVI	600
Dd	538	yAdDdGFATIPgeKPEITIKKAKAEfLYINsKLRPLLELEEFYLRGfFvAKKRRAVI	597
QY	601	DEEGRTIRGLJEVVRBDMSEJAKETOKAVLEALIKESDSEKAVATYKDVVEETAKYOVPL	660
Dd	598	dEEgRTIRGLJEvVRBDMSEJAKetOKAVLEALIKESDSEKAVATyKDVVEETAKYOVPL	657
QY	661	EKLVIHEDITKDLSEYKAGHVAIAKRLAAKIGIKVRGTTISYIVLRGSGKISDRYILL	720
Dd	658	eKlViHeGItKdLSEyKAGhVAIAKRLAAKIGIKVRgTTISyIVLRGsgKISDRYILL	717
QY	721	SEYDRKHXKXUDDYIEMQVYPAVLRLIEAGYKKEDEKCAKQSSQVGLMDMK 773	
Dd	718	seYDRKHXKXUdDYIemQVYPAVLRLIEAGYKKEDEKCAKQSSQVGLMDMK 770	

RESULT	13
AAW48458	
ID	AAW48458 standard; protein; 774 AA.
XX	
AC	AAW48458;
XX	
DT	15-JUL-1998 (first entry)
XX	
DE	Mutant KOD DNA polymerase (144TV).
XX	
KW	KOD DNA polymerase; 3'-5'exonuclease activity; amplification.
XX	
OS	Pyrococcus sp.
OS	Synthetic.
XX	
FH	Key
FH	Misc-difference
FT	144
FT	/note="Changed from Thr in wild-type
FT	to Val in mutant"
FT	
FT	Domain
FT	
FT	137..146
FT	/note="Exo 1 region"
FT	206..222
FT	/note="Exo 2 region"

```

FT      Domain      308..318
FT      /note="Exo 3 region"
XX      EP822256-A2.
XX
XX      04-FEB-1998.
PD      24-JUL-1997; 97EP-0112760.
XX
XX      30-JUL-1996; 96JP-0200446.
PR      29-JUL-1996; 96JP-0198911.
XX
XX      (TOYM ) TOYO BOSEKI KK.
XX      Imanaka T, Kaminura H, Kawakami B, Kawamura Y, Kitabayashi M;
PI      Komatsubarah, Takagi M;
XX
DR      WPI: 1998-102909/10.
XX
XX      Modified thermostable DNA polymerase - with reduced 3'-5'
PT      exonuclease activity
XX
XX      Claim 44; Page -: 54pp; English.
XX
XX      The invention claims a mutated thermostable KOD DNA polymerase of the
CC      present sequence. It has 48% of the 3'-5' exonuclease activity when
CC      compared with the wild-type KOD DNA polymerase. It also has a DNA
CC      extension rate of at least 120 bases/sec, while retaining at least 60%
CC      of its activity after heating at 95 deg. C for 6 hours. The invention
CC      also claims other mutated versions of the polymerase (AAW48449-48461).
CC      Aspartate 141 and glutamate 143 are known to be essential for the
CC      exonuclease activity. The invention shows that efficiency in DNA
CC      amplification can be achieved by using a mixture of two different DNA
CC      polymerases which are almost identical to each other with respect to
CC      thermostability and DNA extension rate but are different in their 3'-5'
CC      exonuclease activity. The mixture may contain a mutant polymerase
CC      (AAW48449-48453) having 0-5% of the 3'-5' exonuclease activity of a
CC      native KOD polymerase used with a second polymerase which may be the
CC      wild-type KOD polymerase or the present one.
CC      N.B. The present sequence is not given in the specification, but is made
CC      up using the given wild-type KOD polymerase sequence and the information
CC      provided in the claims section.
XX
XX      Sequence 774 AA.
SO

```

```

Query Match      80.0%; Score 3220.5; DB 19; Length 774;
Best Local Similarity 77.4%; Pred. No. 6.3e-243;
Matches 598; Conservative 82; Mismatches 90; Indels 3; Gaps 2;

```

```

OY      361 RYAEENNELAPKKPDEEYRRRLRTTYLGQYKKEPERGLMENTITYLDPERCLPSIIVTHN 420
DB      359 rkayerneIapkhpkdekIarr-rqsyeggyvkeperglwentiyldfstlpsliltm 417
OY      421 VSPDITLEREGCKNYDVAIPVIGKFCDFPFPISITIGELITRQELKKMKATIDPIEK 480
DB      418 vspdtlnregckeydvpqgnhfrfckdipgflpsllgdlleerqklkkmkatldpIerK 477
OY      481 MLDYROAVKLANSYGYGMYGPKARWYSKEAESVTAGRHYIEMTIEIEKRGFKVL 540
DB      478 lldyrgaIkIlanSYgygygJararawykceasvtagregylmtlkeekygfKvi 537
OY      541 YADTDFGYATIPGKEPRTIKRAKEFLKYINSKLPGLLELEGYFNGFPAKKRYAVI 600
DB      538 yadtgffatipgdaeetvkkameflnylnaklpaleleyegfykrgffvtkkkyavi 597
OY      601 DEGRITTTGLEVRRDSEIAKETQAKYLEALNKEDSVKAEVLYKDVVEIATKYQVPL 660
DB      598 deegrklttrglelvrrdwseIaketqarvleallkdqdvkavrlvkevtelkskyevPP 657
OY      661 EKLVIHEOITKDLSEYKAIGPVAIAKRLAAGIKVRPCTIISYIVLRSGKISDRVILL 720
DB      658 ekIvIheqIttrldkYkatgphvavakrlaargvklrpgvtvisylvksgrtIdralpf 717
OY      721 SEYDPKHKHYDDPYIENOVLPVAILFAFSYRKREDLKYOSSKOGLDAMWK 773
DB      718 deIdprkKhkydaeyIenqvlpaverllrafykrredlryktrqygsawIk 770

RESULT 14
AAW48450
ID      AAW48450 standard; protein; 774 AA.
XX
XX      AAW48450;
AC      15-JUL-1998 (first entry)
XX
XX      Mutant KOD DNA polymerase (143BA).
DE
XX      KOD DNA polymerase; 3'-5' exonuclease activity; amplification.
XX
XX      Pyrococcus sp.
OS
XX      Synthetic.
XX
XX      Key
FH      Location/Qualifiers
FT      Misc-difference 143
FT      /note="Changed from Glu in wild-type
FT      to Ala in mutant"
FT      Domain
FT      137..146
FT      /note="Exo 1 region"
FT      Domain
FT      206..222
FT      /note="Exo 2 region"
FT      Domain
FT      308..318
FT      /note="Exo 3 region"
XX
XX      EP822256-A2.
XX
XX      04-FEB-1998.
XX
XX      24-JUL-1997; 97EP-0112760.
XX
XX      30-JUL-1996; 96JP-0200446.
PR      29-JUL-1996; 96JP-0198911.
XX
XX      (TOYM ) TOYO BOSEKI KK.
XX      Imanaka T, Kaminura H, Kawakami B, Kawamura Y, Kitabayashi M;
PI      Komatsubarah, Takagi M;
XX
XX      WPI: 1998-102909/10.
XX
XX      Modified thermostable DNA polymerase - with reduced 3'-5'
PT

```


XX Sequence 774 AA:

Query Match 79.9% Score 3218.5; DB 19; Length 774;
 Best Local Similarity 77.4%; Pred. No. 9e-243;
 Matches 598; Conservative 82; Mismatches 90; Indels 3; Gaps 2;

```

QY 1 MIEDTDYITKDGKPIIRFKKENGGEFKIELDPHFQPIYVALKDDSAIDIRAKIGERNG 60
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 mldtdytleedgkpvrltkkenggefkleaydrctfeyfyalikddsaieevkkiteerng 60
QY 61 KIVRVYDAVKVKKKFLGRVDEWVKLIFEHPQDVPALRGKIREHPAVIDIEYDIPAKRY 120
  :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 lvvvrvvkvkqkflgrvewklyflhpqdpalrcklrehgavldiyeydipfakry 120
QY 121 LADKGLIPMEGDEELKLMFADIETFYHGEDEFGKEIIMISYADEPEARVITMKNIDLPY 180
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 lldkglvpmegdeeklmfadenctlyhegefaegpilmisyadeegarvltmknvdipy 180
QY 181 VDVSNREMIKRFVOIVREKDPDVLITYGDNFDLPYLKRAEKLGVTLGLGRDKEHPE 240
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 181 vdvstereemikrfvlrvvkekdpvlltyngdnfdafaylkrcekiglnfalggrds--e 238
QY 241 PRIHRGDSFAVEIKGRHFDLPVVRRTINLPTYTLEAVYEAVALGKTSGAEETIAI 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 239 pkiqrmgdrfavevkgrlfdlpyvlrtlnlpylleavyaavfqpkekvaaeiltpa 298
QY 301 WETESMKLAQYSMEDARATYELCKEFPMEAEI LAKIGOSVWDVSRSSTGNLWEMYL 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 299 wetgenleratysmedakvtyelgkelfpmaeqslrlgslwdvstrssgnlwewfll 358
QY 361 RYAYENELAPKPKDEEYERRRLRTTYLGQYKPEPERGLMENITYLDFRCLYPSIIIVHN 420
  ||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 359 rkayenelapkpdekelaar-rqsyeggykpeperglwenlyldfrslypsliilthn 417
QY 421 VSPDTEREGCKNYDAPIVGKFCDFPGFIPSIIGELITMROEIKKKKATIDPIEKK 480
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 418 vspdtlnregckeydvapgvnhrfckdipgflpsllygdllleerqklkkmkatidpierk 477
QY 481 MLDYRQAVKLANSYGYGYGYPKARWYSKCAESVTANGRYIEMTKEIEEKGFKVL 540
  :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 478 lldyqralkllansgygyygararwyckecaesvtangreyimtlkeieekyfkvl 537
QY 541 YADTGGFVATTIGEKEPETIKKAKKEFLKYNKSLPGLELEYEGFYLKGFYVAKKRYAVI 600
  |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 538 yadtgdfatclpgadaetvkkameflnylnaklpgaleleyegfykrgffvltkkkyavl 597
QY 601 DEGRITTRGLVNRDRDSEIAKETOAKVLEALIKEDSVKAEIYKDYVEELAKYOVPL 660
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 598 deegkitcltrglvnrdrdseiakeqarvleallkdgdvekavrlvkeyleklskyevpp 657
QY 661 EKLVIHQITKDLSEYKAIGPHVATAKRLAAGKIVRPCTIISYIVLRSGKISDRVIL 720
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 658 eklvlheqitrlkdylkatgphvavakrlaargvklrpgtlvsylvlysgslgdraipf 717
QY 721 SEYDEPKKKHYDPPYIENOVLPVLRILEAFGYRKEDLKQSSKOYGLDAMLK 773
  |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 718 defdpkpkhydaeylengvlpaverillrafgyrkedlrygktrgygsawl 770

```

Search completed: May 28, 2002, 14:18:46
 Job time: 104 sec

Q9HH98 PRELIMINARY: PRT: 759 AA.
 ID Q9HH98
 AC Q9HH98
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE DNA POLYMERASE (FRAGMENT).
 GN POL.
 OS Pyrococcus sp. (strain ST700).
 OC Archaea: Euryarchaeota: Thermococcales; Thermococcaceae; Pyrococcus.
 NCBI_TaxID=69015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ST700.
 RA Querellou J.J.E., Cambon M.A., Lesongeur F., Barbier G.;
 RT "Thermococcales taxonomy and phylogeny based on the comparative use of
 RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
 RT genes";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
 CC PYROPHOSPHATE + DNA(N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AJ250332; CAC12847.1; -.
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_POL_B.
 DR InterPro: IPR001648; Ribosomal_S18.
 DR Pfam: PF00136; DNA_POL_B; 1.
 DR Pfam: PF03104; DNA_POL_B_exo; 1.
 DR PRINTS: PR00106; DNA_POLB.
 DR SMART: SM00486; POLB; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN_1.
 DR DNA replication; DNA-binding; DNA-directed DNA polymerase.
 FT NON_TER 759 759
 SO SEQUENCE 759 AA; 88431 MM; A8459B6AD8F3AF3 CAC64;

Query Match 76.5%; Score 3081; DB 1; Length 759;
 Best Local Similarity 74.7%; Pred. No. 5.3e-173;
 Matches 569; Conservative 99; Mismatches 90; Indels 4; Gaps 3;

QY 1 MFDYDITKOGKPIIRIFKKEGFEKIELDHPFQPIYIALLKDSADIDEIKATIGERRG 60
 DB 1 MLDDADYITENKPIIRIFKKEGFEKIEYDHPFQPIYIALLKDSADIDVKKITSERRG 60
 QY 61 KIVRVADVAVKVKKFLGRDVEYWKLFEPHODVPALRGKIREHPVVIDEYDIFPAKRY 120
 DB 61 KIVRVADVAVKVKKFLGRDVEYWKLFEPHODVPALRGKIREHPVVIDEYDIFPAKRY 120
 QY 121 LIDKGLIPMEGDEDEKLMAFDIETFYHEGDEFGKEIIMISYADEEGARVITWKNIDLPY 180
 DB 121 LIDKGLIPMEGDEDEKLMAFDIETFYHEGDEFGKEIIMISYADEEGARVITWKNIDLPY 180
 QY 121 LIDKGLIPMEGDEDEKLMAFDIETFYHEGDEFGKEIIMISYADEEGARVITWKNIDLPY 180
 DB 121 LIDKGLIPMEGDEDEKLMAFDIETFYHEGDEFGKEIIMISYADEEGARVITWKNIDLPY 180
 QY 181 VDVVSNBERMITKRFQVIREKDPDLITNGNDFPLYLKRAEKLGYTLILGROKREHE 240
 DB 181 VDVVSNBERMITKRFQVIREKDPDLITNGNDFPLYLKRAEKLGYTLILGROKREHE 240
 QY 241 VEAVVNERMIRKRLVIRIREKDPDLITNGNDFPLYLKRAEKLGYTLILGROKREHE 240
 DB 241 VEAVVNERMIRKRLVIRIREKDPDLITNGNDFPLYLKRAEKLGYTLILGROKREHE 240
 QY 241 PRHMGDSFAVEIKGRIFHDFPVVARTINPTYLEAVYPAVYKTKSGAGAEIAI 300
 DB 241 PRHMGDSFAVEIKGRIFHDFPVVARTINPTYLEAVYPAVYKTKSGAGAEIAI 300
 QY 239 PRMGLGSLAVIEIGRIHDFLFPVIRKIRINPTYLRLVYAIIGKREKREKYPHEIAEA 298
 DB 239 PRMGLGSLAVIEIGRIHDFLFPVIRKIRINPTYLRLVYAIIGKREKREKYPHEIAEA 298
 QY 301 WETESMKKLAQYSMEDARATYELGKEFFPMEALAKLIGOSVWVSRSTGNLVEWYLL 360
 DB 301 WETESMKKLAQYSMEDARATYELGKEFFPMEALAKLIGOSVWVSRSTGNLVEWYLL 360
 QY 361 RYAVRNELAPKPKDEEERARRLTYYLGGYKKEBERGIMENITLDFRCCLYPSIIIVHN 420
 DB 361 RYAVRNELAPKPKDEEERARRLTYYLGGYKKEBERGIMENITLDFRCCLYPSIIIVHN 420
 QY 421 VSPDTEREGCKNYVAVIPVYKFKDPFGFIPISIGELITTRQELTKKMKATPIEIK 480
 DB 421 VSPDTEREGCKNYVAVIPVYKFKDPFGFIPISIGELITTRQELTKKMKATPIEIK 480
 QY 419 VSPDTEREGCKNYVAVIPVYKFKDPFGFIPISIGELITTRQELTKKMKATPIEIK 480
 DB 419 VSPDTEREGCKNYVAVIPVYKFKDPFGFIPISIGELITTRQELTKKMKATPIEIK 480
 QY 481 MLDYRORAVKLIHANSYGYGMYPKARWYKCECAESVTAWGRHYIEMTIKEIEKEGFKVL 540

DB 479 ILDDYRORAVKLIHANSYGYGMYPKARWYKCECAESVTAWGRHYIEMTIKEIEKEGFKVL 537
 QY 541 VADTGTGFATITGCEKP-ETIKKAKKEFLKYINSKLPGLLEEGFYLGFGVAKKRVAV 599
 DB 538 YIDTGTGLVATITGCEKPNWEEKIKRKALEFVNYINSKLPGLLEEGFYLGFGVAKKRVAV 597
 QY 600 IDEEGRTTGTGLVAVRBDSEIAKETQAKVLAILEKDESVAEVAVKRVVEIKAKQYV 659
 DB 598 IDEEGRTTGTGLVAVRBDSEIAKETQAKVLAILEKDESVAEVAVKRVVEIKAKQYV 657
 QY 660 LEKVIYHEQITKDLSEYKAGIPHVAIAKRLAKGIRVPGTIIYVLNGSGKISDRVIL 719
 DB 658 VEKVIYHEQITKDLSEYKAGIPHVAIAKRLAKGIRVPGTIIYVLNGSGKISDRVIL 717
 QY 720 LSEYTPKHKHYDPDYIENQVLPVLRILAEYKREKEDLKY 761
 DB 718 IEEFDGKKRKHDAEYIENQVLPVLRILAEYKREKEDLKY 759

RESULT 4
 Q9P9N1 PRELIMINARY: PRT: 781 AA.
 ID Q9P9N1
 AC Q9P9N1
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE DNA POLYMERASE B3.
 OS Sulfurisphaera ohwakuensis.
 OC Archaea: Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfurisphaera.
 NCBI_TaxID=69656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TA-1;
 RX MEDLINE=20450680; PubMed=10997874.
 RA Iwai T., Kurosawa N., Itoh Y.H., Kimura N., Horinouchi T.;
 RT "Sequence analysis of three family B DNA polymerases from the
 RT thermophilic crenarchaeon Sulfurisphaera ohwakuensis";
 RL DNA Res. 7:243-251(2000).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
 CC PYROPHOSPHATE + DNA(N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AB032376; BAAS3703.1; -.
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_POL_B.
 DR Pfam: PF00136; DNA_POL_B; 1.
 DR Pfam: PF03104; DNA_POL_B_exo; 2.
 DR PRINTS: PR00106; DNA_POLB.
 DR SMART: SM00486; POLB; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 DR DNA replication; DNA-binding; DNA-directed DNA polymerase.
 SO SEQUENCE 781 AA; 90384 MM; 9B2570EAE30C372C CAC64;

Query Match 29.9%; Score 1204.5; DB 1; Length 781;
 Best Local Similarity 36.6%; Pred. No. 1.1e-62;
 Matches 294; Conservative 141; Mismatches 282; Indels 87; Gaps 21;

QY 2 IDTDYITKOGKPIIRIFKKEGFEKIELDHPFQPIYIALLKDSADIDEIK-ATIGRR 58
 DB 7 ILDFSDVVENKPIYIYIWDKEGNRVVLEKFRFYFALVDSDSYNDIEIKREIKLUSK 66
 QY 59 HKIYRVADVAVKVKKFLGRDVEYWKLFEPHODVPALRGKIREHPVVIDEYDIFPAK 118
 DB 67 PYSPTISIDV--EKKIKGSPVKVLEIETVITAVYRVYRDEYAKIKGASVLEADIRFYM 124
 QY 119 RYLDKGLIPM-----EGDE-----ELKMAFDIETFYH 147
 DB 125 RKSIDINLKPFWIWEAEVEIKENNPRVKKYVELKINKLYEDKIPELKVLAFDIEVY-- 182
 QY 148 EGDGEG-----KGEIIMISYADEEGARVITWKNIDLPYDVVSNBERMITKRFVQVIREK 202
 DB 183 --NKYGSNPRBRDPVILIGVMTKEGKQPLADKYD-----DLRAIRFEPINFGVYD 231

11
RN SEQUENCE FROM N.A.
RP STRAIN-DSM2709;
RC MEDLINE-95238290; PubMed-7721707;
RA Uemori T., Ishino Y., Doi H., Kato I.;
RT "The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-
RT like DNA polymerases.";
RL J. Bacteriol. 177:2164-2177(1995).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
CC PYROPHOSPHATE + DNA(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL; D38574; BAA07580.1; -.
DR HSSP; P56689; ITCO.
DR InterPro: IPR002064; DNA_POL_B.
DR Pfam; PF00136; DNA_POL_B; 1.
DR Pfam; PF03104; DNA_POL_B_exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBc; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase;
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 803 AA; 92656 MM; E9C092F26A8D23FA C6C64;

Query Match 29.6%; Score 1190; DB 1; Length 803;
Best Local Similarity 33.7%; Pred. No. 8.2e-62;
Matches 281; Conservative 164; Mismatches 278; Indels 110; Gaps 23;

OY 1 MIFDTY-----ITKDGKPIIRIFKKEGFEKLEDPHFQPIYALIKD--D 45
DB 8 VLLSSSYELGKEPVILMIGITLIDGKRVL-----LDHRRFPYFALLARGYE 55
OY 46 SAIDIKIKIGERHGIYR--VDAVKKKKFGRDVEWVKLIENHPDVALRGKIREH 103
DB 56 DMVEEIAA--STRLSVKKSPILDAKPLDKRFGKPKAVKITTIPESVRYTRAVAKI 113
OY 104 PAVIDIYVDIYPAKRYLIDKGL-----IPME-----GDEE----- 134
DB 114 EGVEDSLDIEFAMRYLIDKRLYFVYVRIPEVDAGNPGFRVDVRYKVGADPEPLDI 173
OY 135 -----LKMAFOLETYTHED--EFGKEIIMISYADEE--EARYITKNIDLPYDVVS 185
DB 174 TRIDLPKRLVAFDIEVYSRRGSPNARDPVIYLSRSEGKRLIEAGHD----- 225
OY 186 NEREMIKFVOYIREKDDVLTLYNGDNFDLYILKRAEKGVTLLGDKRHEPKIHR 245
DB 226 -DRVLRFEVEYVRAFDPIIYGVNSNHFDPYLMERARRLGIGKLDVTR--KGAEPITISV 283
OY 246 MGDSEFAEIKGRIHFDLEPVYARTINLPTTYLLEAVYLEAVLGKTSKLGAE--IAIWE 302
DB 284 YGH--VSVOGRANVDLYDVAEEMPEIKMKTLLEVAEYLVGVKKSESVYIEMWRIPEYWD 340
OY 303 TEESKKIAQVSMEDARATYELGKEFFPEAEFLAKLIGOSVADSRSSGNLVENVILRV 362
DB 341 DEKKRQLERVALDVRATYGLAEKMLPFAIOLSTVTVGPIDOVGAMGFGFLEWYLMKA 400
OY 363 AYERLEALPNKDEEYERRRLTFTYLGGVKPERGLMENTYILDFRCIYPSIITHNVS 422
DB 401 AYDMELVNPKE-----RGESEYKGAVALKPLKGVHENNVLPDSSMTPSIMIKINVG 454
OY 423 PDTL--EREGCKNND--VAPIVGKFKDPFGFTPSILIGELITMRQETKKMKMA--TIDP 476
DB 455 PDIYIDDPSECKYGGCVAPREVGRFRSPGCFKTYLENLTKLRQVKEKMKKEPPDS 514
OY 477 IEEKMLDYRQAVKLHANSYIGMGPYKARWYSKDCAESVTAMGNHYLDMITKEIEKFG 536
DB 515 PEYRLYDEOKALKVLANASYGMGSHARWCKRCAEAVTAMGNLI--LTAIEYARKLG 573
OY 537 FVLVALADGGEVATIGKEPFTIKKAKFELKYINSKLPGILLEEGYVLNGFV--AKK 595
DB 574 LKVTIGDDSLFVYD-----KERVLEIEVEKEIG--FEIKIDKTYKVFTEAKK 624
OY 596 RYAVIDEGRITTRGLEVVRDMSIARETOAKVLEAILKEDSVKAVEIVKDVVEELAK 655

DB 625 RYVGLLEDGRIDIVGEFAVRGDMCELAKEVQKAEIYLVNLTNGVDKALSTIREVIKQRE 684
OY 656 YQVPLEKVIYHEQITKDSSEKAIQPHVALAKRLAKIGKVRPGITISYILRGSGKISD 715
DB 685 GKVPITKLIIKLTSTRIEYEHDPHYMAARMKMEAEVYSPGKVVYIVYKSGSVSS 744
OY 716 RYILLSEVDPKKHYDPDYIENOVLPAPVLRLEAFGRKEDLKYQSSKQVGL 768
DB 745 RAYPFMWDPDS--TIDVNTYIDHQLVPALRLISFTGTEKOLKAAIVQNSL 795

RESULT 7
IDP9M2 PRELIMINARY; PRT: 785 AA.
AC O9P9M2
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FAMILY B DNA POLYMERASE.
GN POLB3
OS Pyrobaculum islandicum.
OC Archaea; Crenarchaeota; Thermoproteales; Thermoproteaceae;
OC Pyrobaculum.
OX NCBI_TaxID=2277;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GEO 3;
RX MEDLINE-20100754; PubMed-10633098;
RA Kaehler M., Antzanlikian G.;
RT "Cloning and Characterization of a Family B DNA Polymerase from the
RT Hyperthermophilic Crenarchaeon Pyrobaculum Islandicum.";
RL J. Bacteriol. 182:655-663(2000)
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
CC PYROPHOSPHATE + DNA(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL; AF195019; AAF27815.1; -.
DR HSSP; P56689; ITCO.
DR InterPro: IPR002064; DNA_POL_B.
DR Pfam; PF00136; DNA_POL_B; 1.
DR Pfam; PF03104; DNA_POL_B_exo; 2.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBc; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
SQ SEQUENCE 785 AA; 89748 MM; 10BB866A8F3730D C6C64;

Query Match 27.9%; Score 1125; DB 1; Length 785;
Best Local Similarity 33.8%; Pred. No. 5.3e-58;
Matches 273; Conservative 156; Mismatches 284; Indels 94; Gaps 23;

OY 4 DTDYTKGCKPIIRIFK--KENGFEKLEDPHFQPIYALIKDSDAIDELKAIGERHGI 62
DB 9 DITYAVGVSVEIRIFIGLISGGERVLLIDRSKFPYFYV---DCAVCEPAALK--TALSRV 63
OY 63 VRVVAVAVKKKFLGRDVEWVKLIENHPDVALGKIREHPAVDIYEDVIFPAKRYLI 122
DB 64 APIDVQIVERRFLDRSKKFLKVIKIPEDVAKLEAAMSIDPVSGVTEADIRFTMRITV 123
OY 123 DKGLIPM-----EG-----DE-----ELKMAFOLETYTHED--EF 152
DB 124 DMGVVPCSMNNAVEBEGGLGIPYVVSQWYCIDGEPFSLKVMAFDIEVYNNENSGSPDR 183
OY 153 KGKEIIMISYA--DEEARVITKNIDLPYDVVSNEREMIKRVRQVIREKDDPVLTLY 209
DB 184 IRDPVYMLAIKINDCHEEVEFASGR-----DDRGVRAVVDIIRSYDPPVIGY 232
OY 210 NGDNFDLYILKRAKRLVTLIGRDKRHEPKIHRMDSFAVETIKGRIHFDLPFYVVRT 269
DB 223 NSNGDMPFLVERAKAVGAPLKVDRLSNPQOSVYGHM-----STVGRANDVLTNIVEEF 287
OY 270 INLPYVTEAVYEAVLIGKTSK---LGAEEIAAIWETESKKIAQVSMEDARATYELG 325

```

Db      288 PEKLTLDVAE-YFGVMREERVLPGHKITEYMDPKRRLKYYLDDVSTGLA 346
QY      326 KEFFEMAEALAKIGOSVMDVSRSTGNLWENYLLRVAERNEELAPKPEEEVRLRLT 385
Db      347 DKLLPFIQSSVSGPLDQVAAASGNRVEMLLRKAYLGEVAPNR-EEREY----- 400
QY      386 TYLGGYKKEPERGLMENTYLDLFRCLYPSIIIVHNSPDTL---ERECKNYDVAPIVG 441
Db      401 PKGATVLEPKPGMEYEDVLDLDFSSMYPINMKYNLSPDTYLEPPEPPEGVNAPEVG 460
QY      442 YKCKXDPFGFIPISLIGELITMRQELKKMKA-TIDPLEKMDLYRQAVALLHNSYGYM 500
Db      461 HRRFRSPGFPQVLSLVELLRKAVREAKKYPDSEFELIDEROALKVMANATGYL 520
QY      501 GYKRAMYSKECAESYATAMGRHYIEMTIEKEEFGFKVYADTDGFYATIPGKPEPIK 560
Db      521 GWGARMYKREVAESYATAFARALKDVIQ-ARRLGIVVYGGTDSLFFYKHKD----- 573
QY      561 KKAKEFLKYINSKLPGLLEVEYEGFYLRGEFV-AKKRAYAIDEGRITTRGLEVVRDWS 619
b      574 --VDKLTIKYEEKYG--IDIKVDKDYAKVLFTEAKKRYAGLLRDGRIDYVGFVRGDS 629
QY      620 EIMKEOAKYLEALKEIDSEKA----VEYKDVVEEIAKYQVPLEKLYHEQITDLSL 675
Db      630 ELAKDVOLRYELILSRDIVERHGVYKIRIETIRLKNYKFNIDLLIWKTLDELDE 689
QY      676 YKAGHPVAIAKRLAKGKIRGTTIISYIVLSGSKISDRV---ILLSEYDKKHKHYD 732
Db      690 YKAYPPHVAHQILKRNHGRVNGKGTIGYIVYVAGGKYSERAPYLILDDI-----KKIDI 745
QY      733 DYIENQVLEPAVLRLIEAFGYRKEDK 759
Db      746 DYIEROIIPALRLIAEIVGKESDLK 772

RESULT 8
Q9NH85 PRELIMINARY; PRT; 195 AA.
AC 09NH85;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE DNA POLYMERASE (FRAGMENT).
GN POL.
OS Thermococcus chiltonophagus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=54262;
RN [1]
PE SEQUENCE FROM N.A.
NA Querellou J.J.E., Cambon M.A., Lesongeur F., Barbier G.;
RT "Thermococcales taxonomy and phylogeny based on the comparative use of
RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
RT genes.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: A250334; CAC12848.1; -.
DR HSSP: P56689; ITGO.
DR InterPro: IPR02064; DNA_pol_B.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
FT NON_TER 1
FT TER 195
SQ SEQUENCE 195 AA; 22453 MW; 23FA457E4EFA0418 CRC64;

```

Query Match 19.2%; Score 774; DB 1; Length 195;
 Best Local Similarity 73.6%; Pred. No. 3.2e-38;
 Matches 15; Conservative 27; Mismatches 23; Indels 2; Gaps 1;

```

QY      122 IDKGLPMGEDELKLMARDIETFYHGEDEFGGELIMISYADEBARVITMKNIDLPYV 181
Db      1 IDKGLPMGEDELKLMARDIETFYHGEDEFGGELIMISYADEEGKVITMKNIDLPYV 60
QY      182 DVVSNREMIKRPVQIVREKDPVLITTYNGDNFDLPYLIRAKELGVTLLGRDKEHP 241

```

```

Db      61 EVVSSEREMIKRFLKRIKREKDPVITTYNGDSFDLPYLKRAERLGIKPLGRDGS--EP 118
QY      242 KIRMGDSFAVELKGRHFDLPVVRTINLPYTLAEVYEAVALGKTSKLGAEETAAIV 301
Db      119 KMORLGMVAELTKGRHFDLPVVRTINLPYTLAEVYEAVALGKTSKLGAEETAAIV 178
QY      302 ETEESMKKLAQYSMEDA 318
Db      179 ETGGLERVARYSMEDA 195

RESULT 9
Q9NAH1 PRELIMINARY; PRT; 1443 AA.
AC 09NAH1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Y47D3A.29 PROTEIN.
DE Y47D3A.29.
GN Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
PE SEQUENCE FROM N.A.
NA Matthews L.;
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998)
CC -I- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -> N
CC PYROPHOSPHATE + DNA(N).
CC -I- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
CC DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: AL117202; CAB97239.1; -.
DR InterPro: IPR02064; DNA_pol_B.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLBc.1.
DR PROSITE: PS00116; DNA_POLYMERASE_B.1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
SQ SEQUENCE 1443 AA; 163380 MW; EF4EB2E90A50BB8 CRC64;

```

Query Match 15.7%; Score 631.5; DB 5; Length 1443;
 Best Local Similarity 26.0%; Pred. No. 1.3e-28;
 Matches 206; Conservative 155; Mismatches 301; Indels 131; Gaps 27;

```

QY      63 VRYVDAVKKKFLGRDVEYWKLFEPHDPDVPALRGKIREHPAVIDIYEYDIPFAR---- 118
Db      462 VPLERLILIEKKPMG-----I-----I-----I-----I----- 495
QY      119 -----RYLIDKGLPMGEDELKLMARDIETFYHGEDEFGGELIMISY-- 162
Db      496 KYEEFVDEKMRNIKYLESSDAPP---PTIKLLALNVTTLLDRKD---NETCMISMLF 548
QY      163 -----ADEEAR-----VITMKNIDLPY-----VDVSNREMIKRP 194
Db      549 NKPCDLAHSBGSDKLPKCVIYKPPGSGSPYIQRLKETLSKRVKSVANAKALLTLF 608
QY      195 VQIVREKDPVLITTYNGDNFDLPYLIRAKELGVTLLGRDKEHPKIRMGDSFAVE- 253
Db      609 LAMINDEPDIYGHQ-LSATIALIVSRLEKTLKPMWSRISRLKRSINIGIKIGHSKSGW 667
QY      254 --IKGRINHDLPVAVRTINLPYTLAEVYEAVALGKTSKLGAEETAAIWEESKKKA 311

```


Db 668 ELTAGRMILDSKLAMELVKSFPDLTESSQILGAERREIYANEIPOLYVDSKDLISL 727
 Qy 312 OYSMEDARATYELGKEF--FPMEALAKLIGSVWDVSRSSGNLYE--WILLRATYARN 367
 Db 728 NWSMNDLLSLRIYVRLAFLYIOISIVGFI--ASRTIMGGAERNEYILLHAFEKA 784
 Qy 368 EL-APNKPDEEYRRRL-----RTYLGVEYKEPERGIMEN--ITYLDFRCLYP 413
 Db 785 DLIAADKYNSSSEKKKKKEESQEGDEKSKQAQYSGGLVEPKGLYELILLDLSNLSLP 844
 Qy 414 SIYTHANSPOTLRECKKNVDVPIYGVCKPCKPFPPIPSILGELITMROEIKKKMAT 473
 Db 845 SIIQYKINCYTLLYS--KDSDEQLSVPOS--TDIEGVLPREIRKLVCEGRDVKSLMKSE 900
 Qy 474 IDPIKKNLADYRORAVKHLANSYGYMGYPKARWYSKCAESVTMGSHYEMTKEEE 533
 Db 901 RREAKKKOMDROMALKLITANSMTGLGFQYSRFAKPLALVITAKGHE--ILMSKDLVE 959
 Qy 534 KFGKVLVADTDGFYATIPGEKPEIKRKAKEFLKYINSKLPGLLEVEYGFYVGFVA 593
 Db 960 KMGYSVYGDPTDSIMINTNSVDLVAKKLTGNEIKKAVN--KCHRLLELDLDGVEKMLLK 1018
 Qy 594 KRRYA--VIDEGRIT---KGLVYVRDMSIENAKTOAKVLEAIL-----KEDSVKAV 643
 Db 1019 KKKYALVIMPTKVEAKELGLDVRDWSQAKETGAVVADKILDSLRDEWVSSID 1078
 Qy 644 EIVKDVVEIKAYOYVLEKLVITHEOITKOLSEYKAGP--HVAIKRLAANG--IKVRSGT 700
 Db 1079 DLREIRAKLDSGYVPLEMFOISKQLTNRPEQYADYKQCHAAVAQRLNKSKEFLRND 1138
 Qy 701 IISYIVLNG--SGKISDVILISEY--DPKHKHYPDYIYENOVLPVAILTEAFGRKED 757
 Db 1139 IVEYVICEDGTNDPATOQARHTEMAENSALKIDLLYLAQIHPVSRIVEPI---ET 1195
 Qy 758 LKYOSKOVGIDA 770
 Db 1196 DAVRIAENLGLDS 1208

RESULT 10
 Q9HRV9
 AC Q9HRV9 PRELIMINARY; PRT: 901 AA.
 DT 01-MAR-2001 (TRENBLREL, 16, Created)
 DT 01-MAR-2001 (TRENBLREL, 16, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL, 19, Last annotation update)
 DE DNA POLYMERASE B1.
 RN POLB1 OR VNG0521G.
 NS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacterium.
 CC NCB1_TaxID=64091;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA NG W.V., Kennedy S.P., Mahaites G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sirogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenberger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freltas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
 RT "Genome sequence of Halobacterium species NRC-1".
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
 CC PROPHOSPHATE + DNA(N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AE005004; AAG19049.1;
 DR InterPro: IPR002064; DNA_pol_B.
 DR Pfam: PF00136; DNA_pol_B; 2.
 DR Pfam: PF03104; DNA_pol_B-exo; 1.
 DR PRINTS: PR00106; DNAPOLB.

DR SMART: SM00486; POLB; 1.
 DR PROSITE: PS00116; DNA POLYMERASE B; 1.
 KW Complete proteome, DNA replication; DNA-binding;
 KW DNA-directed protein, polymerase.
 SQ SEQUENCE 901 AA; 101987 MW; 2134704B7AB5E49D CRC64;

Query Match 15.5%; Score 626; DB 17; Length 901.
 Best Local Similarity 26.7%; Pred. No. 1.4e-26;
 Matches 242; Conservative 138; Mismatches 319; Indels 208; Gaps 42;

Qy 1 MIFPDYITKGG-----KPIIRIFKKEGFEKIEIDPH-FQPYIT---ALKDSDAIDEI 51
 Db 60 MVTQVDY-TIGSSGDRKRPVLHVFRTDDEAHNRVAGFRYFAPATILSEDLTDVY 118
 Qy 52 -----KAIKGERHGIVRYVDAVKKKFLGADVWKILFHPDVPALCKIREH 103
 Db 119 ITGEDGESIRGEALTKI-----FGRT-----PRDV-----GNIRD- 150
 Qy 104 PAVIDYFYDIPPAKRYLIDK-----GLIPMGDE-----ELKMAF 140
 Db 151 --APDHYADLILFPKRLIDKDITDGVVPPARADGALVYHHDITACEVAADLRVTF 208
 Qy 141 DIETFYHGC-DEFGKELITMISYADEEARVITW-----KNIDLPY----- 180
 Db 209 DIEVDRNGFPEDGEPVCLTSHDSYRDEYVAMLYAPDAPVAPGYDPLTDAD 268
 Qy 181 --VDVVSNEREMIKRFVQIVREKDPDVLITNGDNFDLPYLITKRAEKIGVILLGRKEH 238
 Db 269 IDVRVFDTEAMHDAFLSYIEDTNPDLTGWNFFDPAIDRLDELDPR---TDHDL 324
 Qy 239 PEKPIHRMGDSF-----AVEIKGRHFDLPFVVRRT--INLPYLEAVYEAVIGTKSK 291
 Db 325 DSDRLSRVSEVWTSGMGSPNWKGRVFDLLAYQRTKSELDSYLDVAVGDEIGVCKER 384
 Qy 292 LGAEIIAIIWETESMKRLAOYSMEDARATYELGKE--FPMEALAKLIGSVWDVSR 349
 Db 385 YPG-DIGDLM--EDDPERLLEYNLNLDVELCVIDRQSIIVAFWIDRARKLVCKLEDA--T 439
 Qy 350 STGNIVYEWYLLRVAVERELAPNK--PDEEYRRLRTTYIGGYKKEPERLMENTITLD 407
 Db 440 TPGDVADMYVLHKA--GFVLPSPKQGDAAEFEE-----GGAFFDITVREVSVD 490
 Qy 408 FRCLYPSIIVHNANSPD-LEREGCKNDVAPIYGVCKKDPFGPIPSILGELITMROEI 466
 Db 491 LKSLTPMSMTVITNASPEKVPDEPDGTITPTPGVHNRKRPDGIIRKAVDELLTEREE- 549
 Qy 467 KKKKATIDPIEKMLDY-----RQAVKHLANSYGYMGYPKARWYSKCAESVTAMGR 521
 Db 550 --KARRDHDEPDESEDERYDROQAAVKVIMNSLYGVGMDRFLYDRAMSAGVSTNR 606
 Qy 522 HYIENTIKEIEKFGFKVLVADTDGFYATIPGE--KPEITKK-----AKE 565
 Db 607 EVIDFT--EQAAEEFGYEVAAGDTSVMLEGDMTKAAIAIESPDIEDHINAAYDEFARE 665
 Qy 566 FLKYINSLKPLILELEYGFLRGFVA--KKRYA--VIDEGR-----ITRGLEVVARD 617
 Db 666 QLHADDHR-----FOIEFKLI--RPFQAGKKRRAIGHVYMKDEKADVDDVITGPEYQSD 720
 Qy 618 WSEIAKETQAVLEALIKEDSEKAVELVQDVVEIAKYOVLEKLVITHEOITRDLSEK 677
 Db 721 IAPITKRQVKQEVLDLVNVEGDVAVEEYVHGVIEEFGCGADLDIGIPGIGKRLDYE 780
 Qy 678 AIGHVAIAKLAKKAGIVRGCTIISYVLGS-----GKI----- 713
 Db 781 -----TDTAQVRGAKYANLLGTNFD-----GSKRPRLYLAKVHDFEWHGHEDEGIDTS 831
 Qy 714 -SDRV-----ILISEYD--PKKHYPDPDYIENOVLPVAILTEAFGRKEDLYQ 761
 Db 832 TSDYRFRFRPDVYICFYAQDIPTFVVDYDKMLDKLPKPIERILEALDLSMDEVK-S 890
 Qy 762 SSKOVGL 768

Db 891 GGTGTG 897

RESULT 11

09JURO PRELIMINARY; PRT: 796 AA.

AC 09JURO; 01-MAR-2001 (Tremblrel, 16, Created)

DT 01-MAR-2001 (Tremblrel, 16, Last sequence update)

DE 01-DEC-2001 (Tremblrel, 19, Last annotation update)

GN DNA POLYMERASE (POLB), LARGE CHAIN RELATED PROTEIN.

OS Thermoplasma acidophilum.

OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;

OC Thermoplasma.

OX NCBI_TaxID=2303;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DSM 1728;

RX MEDLINE=20479972; PubMed=11029001;

RA Kuopp A., Graml W., Santos-Martinez M.-L., Korteke K.K., Volker C., Mewes H.-W., Fishman D., Stocker S., Lupas A.N., Baumeister W.; "The genome sequence of the thermophilic scavenger Thermoplasma acidophilum." Nature 407:508-513(2000).

RT -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE + DNA(N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

CC EMBL: A4445065; CAC12036.1; -

CC Interpro: IPR002064; DNA_pol_B.

CC Interpro: IPR000719; Euk_pkinase.

DR Interpro: IPR000577; FGGT_kin.

DR Pfam: PF03104; DNA_pol_B_exo: 1.

DR PRINTS: PR00106; DNAPOLB.

DR SMART: SM00486; POLBc: 1.

DR PROSITE: PS00116; DNA_POLYMERASE_B_1.

DR PROSITE: PS00933; FGGT_KINASES_1; UNKNOWN_1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW Complete proteome; DNA replication; DNA-binding;

KW DNA-directed DNA polymerase.

SO SEQUENCE 796 AA; 91210 MW; DBCDB0F1463EB929 CRC64;

Query Match 15.5%; Score 623; DB 17; Length 796;

Best Local Similarity 25.7%; Pred. No. 1.7e-28;

Matches 211; Conservative 161; Mismatches 314; Indels 134; Gaps 31;

16 IRF-KKNGEKEIEDPHOPITVALLKDDSAIDELKAIKGRHKIVAVNAVKKK 74

19 VLEFGRTREESVAALYFGFRP-YDVPEDEAY-LKVIQND-----PEFVKMDK 67

75 FL--GRDVEWMLFEHPDVPALRGKIREHVAVIDIYEDIPAKRYLIDGL----- 126

68 RLMIIRKYEYVKIYIRSPKAYPEYREK-----CPREVLADIPFHRIYIDLDLSCVK 122

127 -----IPMGDE-----ELKLMADIEFYHEGDEFGKELTMS 161

123 IIGENSIDRETSFTTDIVIRADRIENVDENPMLKVLSEFVENEINRENEVDYKILVIG 182

162 YADEEERAVITWKNIDLPYVDVSNREMAKRFVQVIREKDPVLTLYNGNDLPLIK 221

183 YSYFSGKVTG-----SLSGEODILRSFVDLIRADDPVITGYNIDGDIPYIK 234

222 RAKLGVTLTLGDRKHEPKIRMGDSFAVEIKGRHLPFVVRTITLPTYLEAVY 281

235 RMDRYGKLEITGDSIP-----RIRNQF-WRVHGRILISDTMSVKRILHPKHESLDIYA 289

282 EAVLGKTSKLGAEIAIAIWEETESMKLAOYSMEDARATYELGKEFFPME-----AELA 336

290 NMLGSGKONIDRLHIEDEM--KKRREVIAYCIKADADLIRFEKLMVNNRLMYSVY 347

337 KLIGSVMDVSNKSTGNLVEMWYLLRAVYERBELAPNKPDEEYRRRLRTTYL-GGYVKEP 395

Db 348 KL---PLDDVANGTSNVYDSILIRADRENIGVPMNOHE-----IKTEELDGGYVHSI 398

QY 396 ERGLMENITVDFRCIPSYITVHNVSPTLEREGCKNDVAVIGYKFC--KDPFGFIP 453

Db 399 GAGLYSNVYLDKSMSPSIKYNVCFITLDPNG-----EILSPNGVRLSPKKKGLIP 454

QY 454 SILGELITMRQELKKKAKATIDPIEKMLDYRQRAVKLHANSYYGYGPKARWSEKA 513

Db 455 RILOELMADREDEKRRMKAKAKTDEREFQDGNALKVLMNTFYGLASSFYFTPHKIG 514

QY 514 ESTVANGRHITMTKEIEKFGFVLYADTDCGFVATITGEKEPTIKKAKEFLKTYNK 573

Db 515 SAITFAARETIKCIITLESAN-YRVYIGPTDSVFVESGASADAIKRRKDLSERL-SR 572

QY 574 LPGL---LELEYGFYLRGFEVAKKRA-----VIDEGRITRGLVEYVRDMSIAKET 625

Db 573 EGGITLDFQWVLDPFPSHG---AKKRYAGCVPPDMKGEIVYKGYEVRTDSFDLSQA 629

QY 626 QAKVLEAILKEDSEKAVEIKVDVEEI--AKYQVPLEKLVHEQITKDLSEYKAIGPHV 683

Db 630 LSKVIDFILNRD-IDGAIKYADDLIKKVRAGDPSIDISLVISRTV-KDRQSYRANQESL 687

QY 684 A--IAKRLAAKIKYRPGIISYIVLRGSGKISDRVILLSEDPKKH-----KVDPPY 734

Db 688 ANIRAAKKILDMGETFVPGKYSWIVTNGKR-----TPQEEPYIYGRDLNVKPDWDY 740

QY 735 YIE--NOVLPAVLR-----LEAFGRKEDLK 759

Db 741 YARLSETLGRVLDVFTDLGGOKTASLESFSSSSDEAK 780

RESULT 12

097AH3 PRELIMINARY; PRT: 800 AA.

AC 097AH3; 01-OCT-2001 (Tremblrel, 18, Created)

DT 01-OCT-2001 (Tremblrel, 18, Last sequence update)

DE 01-DEC-2001 (Tremblrel, 19, Last annotation update)

GN DNA POLYMERASE.

DE TVG0859451.

OS Thermoplasma volcanium.

OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;

OC Thermoplasma.

OX NCBI_TaxID=50339;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GS51 / DSM 4299 / JCM 9571;

RX MEDLINE=20570466; PubMed=11121031;

RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S., Kawashima-Ohta Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T., Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.; "Archaeal adaptation to higher temperatures revealed by genomic sequence of Thermoplasma volcanium." Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).

RT EMBL: AP000994; BAB59979.1; -

DR Interpro: IPR002064; DNA_pol_B.

DR Interpro: IPR000719; Euk_pkinase.

DR Interpro: IPR000577; FGGT_kin.

DR Pfam: PF00136; DNA_pol_B_1.

DR Pfam: PF03104; DNA_pol_B_exo: 1.

DR PRINTS: PR00106; DNAPOLB.

DR SMART: SM00486; POLBc: 1.

DR PROSITE: PS00116; DNA_POLYMERASE_B_1.

DR PROSITE: PS00933; FGGT_KINASES_1; UNKNOWN_1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW Complete proteome.

SO SEQUENCE 800 AA; 91833 MW; EA9640C12B134929 CRC64;

Query Match 15.5%; Score 622.5; DB 17; Length 800;

Best Local Similarity 25.9%; Pred. No. 1.9e-28;

Matches 215; Conservative 165; Mismatches 306; Indels 145; Gaps 33;

QY 9 TKGKPIIRIRKENGEEKIELDPHF-----QPIYALLKDDSAI---DEIKAIKGEH 59
DB 27 SREGSVSALE-----FGFKPFDVPEPNEILSSIRNDEVEYKEDKLVHNGSMH 78
QY 60 KRIEVDVAVKVKKFLGHDEVEWMLIFEHPDVPALRGKIREHPAVIDIYEYDIPFAK 119
DB 79 -----NVKRIYIRSPKWKVEYR---RRCF--FEVLAADIPFHR 112
QY 120 YLIDKGL---IMEGDE-----ELKMAPIETFFYHEG 149
DB 113 FYLDPLGACVRIEEDISGTEGNETDVLKIDKIENIPDENVMKVLSPFVENEIRNE 172
QY 150 DEFGKEIMISYADEEERAVITMKNIDLPYDVVSNEREMIKRFOVJIREDDPVLTGY 209
DB 173 NVEDGKILVIGYS-----IMEGDIQ-KKEIGHDEKEILYRIDILIRADDPVITGY 224
QY 210 NGDNFDLPYLKRAKGLVTLGLGRDKEHPEPKIRHMGDSFAVEIKGRIFHPLFPVRRY 269
DB 225 NIDGYDMLVIGRMEQYGIHNLGRGVSYP---RIMDQF-WRVHGRILSDTMNNVKKI 279
QY 270 INLPYITLAEVAVENLGKTKSLGAEELTAIETEESSMKKLAQYSMEDARATYELGKEPF 329
DB 280 IHPKHESLDYIAMKLLGEGKDSIDRLNIEAEWQKROD--EVIISYIKADLTLRIFEKLR 337
QY 330 PME-----AELAKLIGOSVADYSSSTGNLVEWYLLRYAVERNELAPN-----KPEDEEY 379
DB 338 VLERLMKSTYVKL---PLDDVANAGTSTYVDSILIRADRENIGVPMANHLKDELD 393
QY 360 RRRLTYTIGYVYKPERKGLMENTYLDPRCLYPSIITVHNVSPTLREGECKNYDAPI 439
DB 394 -----GGYVHTIGAGLSNVIVLDKFSKMYPSMIITKYVNCFTTLDPKG---EIVSP 440
QY 440 VGYKFC--KDFGFIPTSLIGELITMROELKKMKATIDPEIKKMDYQRAVKLANSTY 497
DB 441 TGIIRLSPKKKKGLPIRLIQLMADRDVYKRMKKAESDEFLYDGIOMAKVLMNFY 500
QY 498 GYMGYPKARWYSKECAESYTANGRHYIEMTIKEIEKFGFKVLVADTGFYATIGEKPE 557
DB 501 GYLAASFYRFTDPKIGSATATAFARETIKHI--DYLESSHRITYGDTDSVYFESGDVNOE 559
QY 558 TIKKAKEFLKYINSKLPLG--LELEYEGFLRGFEVAKKRYA---VIDEGRIITR 609
DB 560 DAIKTKELISQKL-SEEBGLTLDFOVLDPFPSHG--AKKRYAKGCYVPEDMKGEIITK 615
QY 610 GLEVVRRPOMSEIAKETQAKVLEAILKEDSVKAVEIVDVEET--AKYQVLEKLVTHE 667
DB 616 GYEVRTDSFDLSTALSKYIDFILLDRD-VOGAINYADLVKRVNKGDPISIDIESLVISR 674
QY 668 OTTKLSEYKAIGPVA--IAKRLAKGIKYRPGTIISYIYLRSGKISDRVILLSEYD 724
DB 675 TV-KPFSSKANTDSLANTRAAKRLIERGETYVPGMKVSWIYT--NGKKTPEVEPEYIYG 731
QY 735 PK-KKHYDPDYIENQVLPVAILLEAFGYRREDELYOSSKQVGLDAMLWK 774
DB 732 SKLESKPDMDYAK-RLSETLNRVIDVF--RKD--MVOGNRISLSDSFESK 777
RESULT 13
QY 0971C7 PRELIMINARY: PRT: 876 AA.
AC 0971C7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE DNA-DIRECTED DNA POLYMERASE.
GN ST1426.
OS Sulfolobus tokodaii.
OC Archaea: Crenarchaeota: Sulfolobales: Sulfolobaceae; Sulfolobus.
QY NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;

RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamaya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.,
RT "Complete genome sequence of an aerobic thermocacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.",
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000986; BAB66493.1;
KW DNA-directed DNA polymerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 876 AA; 101584 MW; C013A0FB0831F96 CRC64;
Query Match 15.3%; Score 616.5; DB 17; Length 876;
Best local Similarity 26.6%; Pred. No. 4.8e-28;
Matches 231; Conservative 140; Mismatches 316; Indels 181; Gaps 35;
QY 2 IFDPTYIRKDGPIIRIRKENGEEKIELD-PHPQPIYALLKDDSAIDEIKAIKGERHG 60
DB 53 LLOYDYGKSKAVKCLDKETKTIYLDNTGKPRF-----LTDIDPEKVN 100
QY 61 KIVRV-----DAVKKKKFLGRDVEWMLIFEHPDVPALRGKIREHPAVIDIYE 111
DB 101 KIPKVVDRPSPDLHETVYKI-DPYSGNKIKLTKIYVKNPDLAVRRMRNSVPR-----AYE 153
QY 112 YDIPFAKRYLDKGLP-----MEGDE----- 133
DB 154 AHIKYFNMYIDGLIPLPYVKKKGLDVOYRELKGEVEIRKAFADSEMTEKAVND 213
QY 134 -----ELKMAPIETFFY-----HEGDFEGKEITMSYA-DEBEARVITWK 174
DB 214 WPIRESEVPYKRAVIDIEVYTPKGRIPDEKAEF---PITSISLAGNOSTKRVVLV 270
QY 175 NIDL-----PYVDVSNEREMIKRFOVJIREDDPVLTGYNDNFDLPYLKRAK 226
DB 271 REDVNSQITKHVDIYETKSERELIRREFDIL--LDYPIITFNDDDDIPYIVYRALK 328
QY 227 GVTLLGDKDKEHPEPKIRHMGDSFAVEIKGRIFHPLFP-VVARTI-----NLPTYLE 278
DB 329 NPT-----PEELIPDILNDEGKYLA---GIHIDLYKFFFNRAIRNAYFECKYNEYND 378
QY 279 AVEVAVLKTSKSLAEELTAIETEESSMKKLAQYSMEDARATYEL---CKEFPMEAE 334
DB 379 AVATRLDMSKVKL--DTLISFLD---LDKLIENSRDAELTKLTFFNNNLVWKLLIL 432
QY 335 LAKLIGOSVADYSSSTG---NLVEWYLLRYAVERNELAPNKPDEEYRRRLRTT--- 386
DB 433 LARISKMGLELITREVSQWIKNLVYMEHR--RNNLIPKEELIRSSQIKTAALIK 488
QY 387 ---YLGCVYKPERKGLMENTYLDPRCLYPSIITVHNVSPTLREGECKN---YDAPI 439
DB 489 GKRYGAVVDDPPAGVFNVVNVLDFASLYPIIRMMNISLEVVDENCKNEYVEDTGE 548
QY 440 VGYKFCDFGFIPTSLIGELITMROEI-KKKKA-TIPIEKMDYQORAVKLANSTY 497
DB 549 VLIHTCKDKPGITAVITGLIRDFRKYVYKKAASSONISEQGSYDVQVRAKVIINATY 608
QY 498 GYMGYPKARWYSKECAESYTANGRHYIEMTIKEIEKFGFKVLVADTGFYATIGEKPE 557
DB 609 GYFGAENFPPLVAPAAVESYTAIGRVITTYVYC-RSIGLOLVYDGTSMFLMNS----- 663
QY 558 TIKKAKEFLKYINSKLGLLELEYEGFLRGFPVA---KKRAVIDIEGKITRGLEV 613
DB 664 --KEKLEIKIFVKGKFG--LDLEVDKIYK--FVAFSGLKNNYGYVPDGTDTJKGMA 716
QY 614 VRDMSIAKETQAKVLEAILKEDSVKAVEI-----VKDVEIKKYQVPLEKVIH 666
DB 717 KKRNTPEITKKFENYKQVLTITNSPDPIKRDLEKIKIKIYKLVKLNKGYNDLELFR 776
QY 667 EOTTKLSEYKAIGP-HVAIAKRLAKGIKYRPGTIISYIYLRSGKISD-RVILLSEYD 724

Db 777 VMLSKPLESTKNTPOHVKALOLRSGVNVLPDRIIMFVKVSKDGVPOLAKLSE-- 834
Qy 725 PKRHKYDPDYIENOVLPVLRILEAFG 752
Db 835 -----IDVDKYID-AVRSSTFEQILKAFG 856

RESULT 14
Q91VT0
ID Q91VT0 PRELIMINARY; PRT: 1105 AA.
AC Q91VT0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE DNA POLYMERASE DELTA 1, CATALYTIC DOMAIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA TISSUE-BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009126; AAH09128.1;
SQ SEQUENCE 1105 AA; 123744 MW; 20323690DDA72406 CRC64;

Query Match 15.3%; Score 616.5; DB 11; Length 1105;
Best Local Similarity 27.2%; Pred No. 6 6e-28;
Matches 206; Conservative 111; Mismatches 297; Indels 123; Gaps 27;

Qy 60 GKIVRVDAVKKKFLGRDVEY--WKLFEPH-----QDVPLRGKIREHPAVIDIYE 111
Db 269 GKIVR-----RAEKKATLCQLEVDVLSMDVISHPREGQORIPLR-----VLS 312
Qy 112 YDIPFAKRYIDKGLIP-MEGDEELKIMADIEFYHEGDEFGKIGELIMISYADEEAR 170
Db 313 FDEECAGR-----GIFPERPDVIOICSLGR--WGEPEPLRLALTLPAPILGAKY 366
Qy 171 ITWKNDLPYVDVNSEREMIKRFOIVREKDPDLITYGDNFDPYLIRAEKIGVTL 230
Db 367 QSE-----KEEDLLQAMADFILAMPDBDVTIGNIQNFDPYLISRAQILKVD 415
Qy 231 --LLG-----RKHEPERKIHMGDSFAVEIKGRIFHDLFPVVRTINLPYTLEA 279
Db 416 FPELGVHTGLRSNIRDSFOSRQVGR--DSKVISMVGRVQMDLQVLLREHKLRSTLNA 474
Qy 280 YVEAVLGKTKSKGAEIATAIWE--TEESMKKLAQYSMEDARATYELGKEFFM----- 331
Db 475 VSRHFLGEQKEDVQHSITDLONGNEQTRRLRAVYCLKDA-----FLPLRLLERLM 525
Qy 332 ---EAEKALIGOSV-WDVSRSSTGNLVEWYLLRVAVERNELAPNKPDE--EEYRRRLR 384
Db 526 VLVNNVEMARVTVGPLGYLLTRGOQKVV--QLLRQAMRGMLMPVYKREGSD----- 578
Qy 385 TTYLGTVKEPEBGLME-NITYIDFRCLYPSIIYTHNVSPDLERBGC-----KNYDV 436
Db 579 --VTGATVIEPLKGYDVPATLDFSSLPSIMAHNLCTTLRLRGAQKLGKIDDEF 636
Qy 437 APIVKKFCDF--PGFISILGELTMRQELIKKKMATIDPIEKKMLDYRQAVLHAN 494
Db 637 KPTGGEFVKSSVRKGLPOLLENLSARRAVAELAQETDPLRQVLDGRQLALVLSN 696
Qy 495 SYGYNGYPAKRWYKSCAESVTAMGRHYIEMTIKEIEKEF-----GF-----KVLVADTGG 546
Db 697 SYVGFIGAOGKRLPCLFIEQSOTVGFROMIEKTKOLVESYTYVENSYDANAKVYGDJTS 756
Qy 547 FYATIGEPETIKKAKKFLKYNKLPOLLELEVEGYLRFVAKKRYAVI----- 600
Db 757 VMCRFGVSSVAEAMSLGREANMVSSHPSPIRLEFEKVFYPLLISKRKYAGLFRSSS 816
Qy 601 DEGRITTKGLEVVRDMSIAKETQAKVLEALKEDSEKAVEIKVDVEETIAKIQVPL 660

Db 817 DAHDKMDCKGLEAVRRDNCPLVANLVTSRLRLVYRDPGAAVAHAKVDISLCKNRIDI 876
Qy 661 EKVIHEOITKDLSEYKAIQPAIKRLAKGIVRP--GTIISYIVRGSG-----K 712
Db 877 SOLVITKELTRAADYAGKOAHELAEKRRKPPGSAFSLGDNVPYVIIGAKGVAAK 936

Qy 713 ISDRVILLSEYDPKHKYDPDYIENOVLPVLRILE 749
Db 937 SEDPLFVLEHSLP-----IDTQYVLEQOLAKPLRLRIFE 969

RESULT 15
Q96H98
ID Q96H98 PRELIMINARY; PRT: 1107 AA.
AC Q96H98;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE POLYMERASE (DNA DIRECTED), DELTA 1, CATALYTIC SUBUNIT
DE (125KD).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA TISSUE-LYMPH, AND LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008800; AAH08800.1;
SQ SEQUENCE 1107 AA; 123611 MW; 10A8A3271916EDBA CRC64;

Query Match 15.3%; Score 616; DB 4; Length 1107;
Best Local Similarity 27.5%; Pred No. 7 1e-28;
Matches 202; Conservative 134; Mismatches 309; Indels 90; Gaps 24;

Qy 68 AVKVKKFLGRDVE--VWKLFEHPDVPALRGKIREHPAVIDIYEDIPFAKRYLIDK 124
Db 274 ALTKREKATQCLEADVLSMDVSHPREGQORIPLR-----IAPLRVLSFDEECAGR-----K 323
Qy 125 GLIP-MEGDEELKIMADIEFYHEGDEFGKIGELIMISYADEEARVITWKNDLPYVDV 183
Db 324 GIFPERPDVIOICSLGR--WGEPEPLRLALTLPAPILGAKVQSYE----- 372
Qy 184 VNEREMIKRFOIVREKDPDLITYGDNFDPYLIRAEKIGVTL--LLG----- 233
Db 373 --KEEDLLQAMSFITRIMPDPVTIGNIQNFDPYLISRAQILKVDYTFPFLRGVAGLCSN 430
Qy 234 -RKHEPERKIHMGDSFAVEIKGRIFHDLFPVVRTINLPYTLEAVYEAVALGKTKSKL 292
Db 431 IRDSSFOSQOTGR--DTKVISMVGRVQMDLQVLLREYKLRSTLNAVSFHFLEGEQEDV 489
Qy 293 GAEIATAIWE--TEESMKKLAQYSMEDARATYELGKEFFM--EAEKALIGOSV-WDVS 348
Db 490 QHSITDLONGNEQTRRLRAVYCLKDAYLPLRLLELMVLVNAVEMARVTVGYLLSLR 549
Qy 349 SSTGNLVEWYLLRVAVERNELAPNKPDE--EEYRRRLRTTYLGQYKEPEBGLME-NITY 405
Db 550 GQOQKVV--QLLRQAMHEDLMPVYKSEGED-----YTGATVIEPLKGYDVPAT 600
Qy 406 LDFRCLYPSIIYTHNVSPDLERBGC-----CKNYDVAIVGKFCDF--PGFISIL 456
Db 601 LDFSSLYPSIMAHNLCTTLRPGTAQKLGLEDGFITTPGDEPVKTSVRKGLLPOL 660
Qy 457 GELITTRQELIKKKMATIDPIEKKMLDYRQAVLKHANSYTYGMYPKRWYKSCAESV 516
Db 661 ENLSARKRAKAEIAKETDPLRQVLDGRQLALKVSANVSYFTGQVQKLPCLFIEQSOTV 720
Qy 517 TMGRHYIEMTIKEIEKEF-----GF-----KVLVADTGGYAIIPGKPEITKKKAEFLK 568
Db 721 TGRGROMIEKTKOLVESYTYVENSYDANAKVYGDJTSVMCRFGVSSVAEAMALGREAD 780

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:18:07 ; Search time 14.97 Seconds
(without alignments)
2001.933 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026

Sequence: 1 MIFPDYITKDGKPIIRPK.....KEDLKYSKQVGLDAWLKK 774

Scoring table: BLOSUM62

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	3446.5	85.6	1829	1 DPOL_THEST
2	3356	83.4	1702	1 DPOL_THELI
3	3253.5	80.8	773	1 DPOL_THES9
4	3204.5	79.6	775	1 DPOL_THES9
5	3164.5	78.6	771	1 DPOL_PYRAB
6	3163.5	78.6	771	1 DPOL_PYRAB
7	3151	78.3	775	1 DPOL_PYRBU
8	2974.5	72.9	1312	1 DPOL_PYRSD
9	2886	71.7	1235	1 DPOL_PYRHO
10	2809	69.8	1523	1 DPOL_PYRHO
11	2782	69.1	1671	1 DPOL_PYRKO
12	2769.5	68.8	1699	1 DPOL_THES8
13	2535.5	63.0	1668	1 DPOL_THES9
14	1392	34.6	824	1 DPOL_METVO
15	1301	32.3	781	1 DPOL_ARCFU
16	1292	32.1	1634	1 DPOL_METVA
17	1209.5	30.0	586	1 DPOL_METTH
18	1138	28.3	784	1 DPOL_AERPE
19	697.5	17.3	1105	1 DPOL_ORYSA
20	693.5	17.2	1088	1 DPOL_ORYSA
21	691	17.2	763	1 DPOL_SORBN
22	676.5	16.8	764	1 DPOL_SULSH
23	676.5	16.8	1081	1 DPOL_SULSO
24	657.5	16.3	1038	1 DPOL_ARATH
25	657.5	16.3	1086	1 DPOL_CANAL
26	624.5	15.5	1092	1 DPOL_DROME
27	621.5	15.4	1097	1 DPOL_DROME
28	619.5	15.4	1103	1 DPOL_YEAST
29	616.5	15.3	872	1 DPOL_SULOH
30	616.5	15.3	882	1 DPOL_SULSO
31	615.5	15.3	1105	1 DPOL_MOUSE
32	615.5	15.3	1103	1 DPOL_MOUSE
33	614.5	15.3	1094	1 DPOL_PLAFK

ALIGNMENTS

RESULT ID	DPOL_THEST	STANDARD:	PRT: 1829 AA.
AC	033845:		
DT	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	DNA polymerase (EC 2.7.7.7).		
GN	POL.		
OS	Thermococcus sp. (strain TY).		
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.		
OX	NCBI_TaxID=110163;		
RN	[1]		
RP	SEQUENCE FROM N-A.		
RP	MEDLINE:0090467; PubMed:9434178;		
RA	Mathias F. Frey B. Nutritional G.;		
RT	Cloning and characterization of a thermostable alpha-DNA polymerase		
RT	from the hyperthermophilic archaeon Thermococcus sp. Ty.;		
RL	Gene 204:153-158(1997).		
CC	-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate + [DNA](N).		
CC	-1- PPM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE THREE INTERFERING REGION		
CC	(INTERNS) BELONGS TO DNA POLYMERASE TYPE-B FAMILY.		
CC	-1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: Y13030; CAA73475.1; -.		
DR	HSP: P56689; JMC.		
DR	InterPro: IPR002064; DNA_POL_B.		
DR	InterPro: IPR003586; Hinc.		
DR	InterPro: IPR003587; Hinc.		
DR	InterPro: IPR002203; Intein.		
DR	InterPro: IPR004042; Intein_endonuc.		
DR	Pfam: PF00116; DNA_POL_B; 4.		
DR	Pfam: PF03104; DNA_POL_B_exo; 1.		
DR	PRINTS: PR00379; INTEIN.		
DR	SMART: SM00305; Hinc; 3.		
DR	SMART: SM00306; Hinc; 3.		
DR	SMART: SM00486; POLB; 1.		
DR	PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.		
DR	PROSITE: PS00881; PROTEIN_SPLICING; 3.		
KW	Transferase; DNA-directed DNA polymerase; DNA replication;		
KW	DNA-binding; Hydrolyase; Endonuclease; Autocatalytic cleavage;		
KW	Protein splicing.		
FT	CHAIN 1	409	DNA POLYMERASE, 1ST PART (POTENTIAL).
FT	CHAIN 410	769	INTEIN I.
FT	CHAIN 770	855	DNA POLYMERASE, 2ND PART (POTENTIAL).
FT	CHAIN 856	1392	INTEIN II.

34	611	15.2	1107	1	DPOL_HUMAN	P28340 homo sapien
35	609	15.1	1513	1	DPOL_OXYTR	027155 oxytricha t
36	605.5	15.0	1106	1	DPOL_BOVIN	P28339 bos taurus
37	587	14.6	1492	1	DPOL_OXINO	Q94636 oxytricha n
38	584	14.5	1451	1	DPOL_RAT	089042 ratius norv
39	583.5	14.5	1462	1	DPOL_HUMAN	P09884 homo sapien
40	581	14.4	1081	1	DPOL_CAEL	P09829 caenorhabd
41	578.5	14.4	875	1	DPOL_SULAC	P5690 sulfolobus
42	570.5	14.2	959	1	DPOL_AERPE	093745 aeropyrum p
43	570	14.2	1465	1	DPOL_MOUSE	P33608 mus musculu
44	569.5	14.1	1243	1	DPOL_ORISA	048653 oryza sativ
45	548.5	13.6	1339	1	DPOL_TRYBB	P27721 trypanosoma

FT CHAIN 1393 1441 DNA POLYMERASE, 3RD PART (POTENTIAL).
 FT CHAIN 1442 1598 INTEIN II.
 FT CHAIN 1599 1829 DNA POLYMERASE, 4TH PART (POTENTIAL).
 SO SEQUENCE 1829 AA; 211875 MW; A113ABBC57EB9CB3 CRC64;

Query Match 85.6%; Score 3446.5; DB 1; Length 1829;
 Best Local Similarity 42.2%; Pred. No. 3.2e-186;
 Matches 771; Conservative 0; Mismatches 3; Indels 1055; Gaps 3;

QY 1 MIFDTYITPKDGPRIIRIFKENGKEFKIELDPHOPITYALLDDSDAIDEIKAIKGRHG 60
 Db 1 MILDYITRKDGPIIRIFKENGKEFKIELDPHOPITYALLDDSDAIDEIKAIKGRHG 60
 QY 61 KIVRVDAVAVKKKFLGRDVEVKLIFEHPDYPALRGKIREHPAVIDIYEYDIPKARY 120
 Db 61 KIVRVDAVAVKKKFLGRDVEVKLIFEHPDYPALRGKIREHPAVIDIYEYDIPKARY 120
 QY 121 LIDKGLIPMEGDELKLMADITETHEGDEFGKGEIIMISYADEBARVITWKNDLPY 180
 Db 121 LIDKGLIPMEGDELKLMADITETHEGDEFGKGEIIMISYADEBARVITWKNDLPY 180
 QY 181 VDVSNEREMIKRFPVOIVREKDPVLITVINGDNFDLPYLIRAKELGVTLIGRDKHPE 240
 Db 181 VDVSNEREMIKRFPVOIVREKDPVLITVINGDNFDLPYLIRAKELGVTLIGRDKHPE 240
 QY 241 PKIHRGDSFAVEIKGRHEDLPVVRRTINLPTYLEAVYEAVALGKTSKGAEEIAAI 300
 Db 241 PKIHRGDSFAVEIKGRHEDLPVVRRTINLPTYLEAVYEAVALGKTSKGAEEIAAI 300
 QY 301 WEFEESMKKLAQYSMEDARTYELGKEFFPMEMELAKLIGOSWDVSRSTGVLVEMYLL 360
 Db 301 WEFEESMKKLAQYSMEDARTYELGKEFFPMEMELAKLIGOSWDVSRSTGVLVEMYLL 360
 QY 361 RVAYERNELAPNKPDEEYRRLRTTYLGIVYKPERGKLENTYDFERC----- 410
 Db 361 RVAYERNELAPNKPDEEYRRLRTTYLGIVYKPERGKLENTYDFERC----- 410
 QY 411 ----- 410
 Db 421 GKIVINSDVKEGDYILIGDMQVKKVMKYHEGKLININGLCTPNHNVPVVTEHDQ 480
 QY 411 ----- 410
 Db 481 TRIRDSIANSFLSGKVKGIITTKLEKIAEFENKPNPSEELIKGELSGIILAEGTLRK 540
 QY 411 ----- 410
 Db 541 DIEFDSRSGKKRISHQYREITIGENKELLERILYIPDKLFGIRSVAKKGDYNAKTI 600
 QY 411 ----- 410
 Db 601 TPAKKAVALQIEELLKNIESLAVAPLRCGFEBDATTNKRSTIVVTGOTNNKMKIDIVA 660
 QY 411 ----- 410
 Db 661 KLLDSGIPYSRYEYKYIENGKELTKHILETTGRDGLILFQTLVGFISSEKNLEKAIE 720
 QY 411 ----- 410
 Db 721 VREMNRLKNNSFYNSTFEVYSSEYKGEVYDLTLEGNPYFANGILTHNSLYSIIVTHN 780
 QY 421 VSPDTLEREGCKNYDVAPIVGYKCKDPGFIPIISIGELITMROEIKKKKATIDPIEKK 480
 Db 781 VSPDTLEREGCKNYDVAPIVGYKCKDPGFIPIISIGELITMROEIKKKKATIDPIEKK 480
 QY 481 MDYRQRAVLAH----- 493
 Db 841 MDYRQRAVLAH----- 493
 QY 494 ----- 493
 Db 901 VDNIFAFSLNKESKSEIKKVKALLIRHKYKGAEYEVLEANSGRKIHITRGHSLFTIRNGKI 960

QY 494 ----- 493
 Db 961 KEIWGEVYKVGDLIIVPKKVKLNEKEAVINIPBLISKLPEDOTADYVMTTPVGRKNQFK 1020
 QY 494 ----- 493
 Db 1021 GMLRTLKWIKEESKRIRTEFNRYLPHLEELGFVKLLPRGYEYTDWEGKLRKRYQLEYKLVK 1080
 QY 494 ----- 493
 Db 1081 NLRNGNKKREYLVRENDIKDSVSCFPKKELEEMKIGTKXGFRKCKILKYDEDFGKFLGY 1140
 QY 494 ----- 493
 Db 1141 VSEGYAGAOKNKTGMSYSVKLYNENPNVLKDKKNIAEKFFGKRVGRKNQVDIPKKMAYL 1200
 QY 494 ----- 493
 Db 1201 LAKSLGVTAENKRIPSIIFDSSEPVMAFLRAYFVGDDIHPSKRLRSTKSELLANQL 1260
 QY 494 ----- 493
 Db 1261 VELLNSLGVSSTIGFDSGYRYVINEDLPFLQTSKQKNTYFNLLPKVELEIFGRKFKQ 1320
 QY 494 ----- 493
 Db 1321 KNITFEKRELADSGKLDKRRVKLIDLNGDIVLDRVKNVEKREYGYVYDLSVEDNEN 1380
 QY 494 ----- 542
 Db 1381 FLVGFLLVAHNSYGYMGYPKARWYKKECAESVTAMGRHYIMTJKEIEKGFVKVLA 1440
 QY 543 D----- 543
 Db 1441 DSVTGDTEIIVKRNHGRHEFPIEKLEPRVDYRIGKEKCYLDEVALITLDNRKILMKKY 1500
 QY 544 ----- 543
 Db 1501 PYVNRRAKKKYYRIMTNSWYIDVTEHDSLVAEDGLKEARPMIEGKSLATKDDLSG 1560
 QY 544 ----- 565
 Db 1561 VEYIKPHAIEISYNGVYDIEVEGTHREFPANGILVHNHDFGFAATIPGEKPETIKKRAE 1620
 QY 566 FLKYINSKLPGLLELEYEGEYLRGFVAKKRYAVIDEGRITTRGLEVYARDSSEIAKET 625
 Db 1621 FLKYINSKLPGLLELEYEGEYLRGFVAKKRYAVIDEGRITTRGLEVYARDSSEIAKET 1680
 QY 626 QAKVLEAILKEDSVKAVEIVKDVVEIAKYOVPLEKLVYHEQITYDLSFYKAIGPHVAI 685
 Db 1681 QAKVLEAILKEDSVKAVEIVKDVVEIAKYOVPLEKLVYHEQITYDLSFYKAIGPHVAI 1740
 QY 686 AKRLAKGTVKRGCTIISYTVLRGSKISDRVILLSEYOPKKKKYPODYIENOVPAVL 745
 Db 1741 AKRLAKGTVKRGCTIISYTVLRGSKISDRVILLSEYOPKKKKYPODYIENOVPAVL 1800
 QY 746 RILEAFGYRKEDLKYOSSKQVGDAMLK 774
 Db 1801 RILEAFGYRKEDLKYOSSKQVGDAMLK 1829

RESULT 2
 DPOL THELI STANDARD: PRT: 1702 AA.
 AC P30317:
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) (Vent DNA polymerase) [Contains:
 DE Endonuclease P1-T111 (EC 3.1.-.-) (T11 pol-1 intein) (IVPS2);
 DE Endonuclease P1-T111 (EC 3.1.-.-) (T11 pol-2 intein) (IVPS1)].
 GN POL.

OY	1	MILDTYITKDGKIIIMFKKENGKIELDPHFQYIYALLKDSADIEIKAIENGHG	60
Db	1	MILDTYITKDGKIIIMFKKENGKIELDPHFQYIYALLKDSADIEIKAIENGHG	60
OY	61	KIVRVDAVYKKKKFLORDVYEWKLLFJEHQDPALRGKIREHPAVIDYENDIPFARY	120
Db	61	KIVRVDAVYKKKKFLORDVYEWKLLFJEHQDPALRGKIREHPAVIDYENDIPFARY	120
OY	121	LIDKGLIPMEGDEELKIMADIEFYHEGDEFGKEIIMISYADEEARVITWKNDIPY	180
Db	121	LIDKGLIPMEGDEELKIMADIEFYHEGDEFGKEIIMISYADEEARVITWKNDIPY	180
OY	181	VDVYNSNEMKRRVVOYVREKDDVITITNGDNFDLPYIKRAEKLGYTLILGRKHE	240
Db	181	VDVYNSNEMKRRVVOYVREKDDVITITNGDNFDLPYIKRAEKLGYTLILGRKHE	240
OY	241	PKIHRMGDFAVEIKGRIHFDLPVYVARTINLPTYLEAVYEAUIGTKRSKGAEIAI	300
Db	241	PKIHRMGDFAVEIKGRIHFDLPVYVARTINLPTYLEAVYEAUIGTKRSKGAEIAI	300
OY	301	WETESKKKLAQYSMEDARITYELGKEFPFMEAEKLLIGOSVMDVSRSSGTGNLWYLL	360
Db	301	WETESKKKLAQYSMEDARITYELGKEFPFMEAEKLLIGOSVMDVSRSSGTGNLWYLL	360
OY	361	RVAERNELAPNKRDEEYRRRLTYTLYGVAEPERGLMENTYLDPCXPSTIYVHN	420
Db	361	RVAARNELAPNKRDEEYRRRLTYTLYGVAEPERGLMENTYLDPRSLPSTIYVHN	420
OY	421	VSPDTEREGCKNYDAPIVGYKFCDDPGFIPISILGELITMKOEIKKMKATDPIEKK	480
Db	421	VSPDTEREGCKNYDAPIVGYKFCDDPGFIPISILGELIAMRODIKKMKSTIDIPIEKK	480
OY	481	MLDYRORAVYKLAH-----	493
Db	481	MLDYRORAIKLAHSILPNWDLPIENGELIKYKICEFTINSYMEKOKENYKTVENTEYLE	540
OY	494	-----	493
Db	541	VNNLFASFNNKIKESIEVKKVVALIRHKYKAYEIQLSGKKNITAGHSLFYRWGEI	600
OY	494	-----	493
Db	601	KEVSGDGIEGDLIVAEKKIKLNEKGV SINIPELISDLSSEETADIVMTISAKGRNPFK	660
OY	494	-----	493
Db	661	GMLRTLRMHGGEENRRIRITRNRYLPHLEKLGILKILPRGYEVTDMERLKKYQLYEKLAG	720
OY	494	-----	493
Db	721	SVXTNGMKRELYVMFNIEIKOFISVFPQKLEBEKIGTLNGPPTNCLIKVDEDFGLGYY	780
OY	494	-----	493
Db	781	VSEGYAGOKNKTGSIYSVKLYNEDPNVLESKNVAEKFGRVYRNCVSISSKMAYL	840
OY	494	-----	493
Db	841	VMKCLCGALENKRTIPSVIITSPERPAMSHLEAYFTGGDDIHPSKRFRSLTKSELLANOL	900
OY	494	-----	493
Db	901	VFLNLSGISSVKIGFDSGYRVYINEDLOFPQTSREKMTYXSNLIPKILBDVGEKFEQ	960
OY	494	-----	493
Db	961	KNNYFKFKELVDGSKLNRREKAKLEEFJINGDIYLDRAVSYEKDYEGYVYDLSVEDNEN	1020
OY	494	-----NSYYGYMGYPRKAWYSKCEAESVTAMGHNHYIEMTIKEIEEEKFGKUYLA	542

```

Db 1021 FLVGFGLLAHNSYGYMGYPKARWYSKECAESVTAMGRHYEMTLREIEKKGFLVLA 1080
QY 543 D----- 543
Db 1081 DVSAGESEIIIRONGKIRPVKKIDLESKYDSTIGEKEYCLEGVEALTLDDDGKLVKPY 1140
QY 544 ----- 543
Db 1141 PYVMRRANKRMFRILWNSWYIDVTEHSLIGLNTSKTKTAKKIGERLKEVKKPELGG 1200
QY 544 ----- 543
Db 1201 AKSLICPNAPLKDEKTKTSELAVKFWEVLVLIVGDGNGGDSRMAEYVGLSTGDAEE 1260
QY 544 ----- 543
Db 1261 IKOKLEPLKTYGVISNYYPKNEKGPNTLAKSLVKFMKRHFKEDEGRKKIPFEMTELV 1320
QY 544 ----- 543
Db 1321 TYIEAFGLRGSADGVTTIRKGVPEIRLNTIDADFLREVRKLLMIVIGISNSIFAETTPNR 1380
QY 544 ----- 543
Db 1381 YNGVSTGYTKHLRINKRMFPAERIGFLIERKOKLLEHLKSAVKRNTIDEGFDLVHWK 1440
QY 544 ----- 543
Db 1441 KVEEIRYEGVYDIEVEETHRFANLIVHNTDGFATIRGEKPELKKKAKFLVYNS 1500
QY 573 KLPGLLELEEGFYLRGFVAKKRYAVIDEGRIITRGLEEVVRDMSLAKETQAKVLEA 632
Db 1501 KLPGLLELEEGFYLRGFVAKKRYAVIDEGRIITRGLEEVVRDMSLAKETQAKVLEA 1560
QY 633 ILKEDSVKAVELVKDVEIKAYOYPLKELVYHEQITDLSYKAGISPHVATLAKLAK 692
Db 1561 ILKESGVEAVEVKDYVERIKARYPLKELVYHEQITDLSYKAGISPHVATLAKLAK 1620
QY 693 GIKVRPGTIISYVLRGSGKISDRVILLSEYDPKKHRYDPDYIENQVLPVLRILAEAG 752
Db 1621 GAKVKKGTIIISYVLRGSGKISDRVILLSEYDPKKHRYDPDYIENQVLPVLRILAEAG 1680
QY 753 YRKEDLKYSKQVGLDANLKK 774
Db 1681 YRKEDLKYSKQVGLDANLKK 1702

RESULT 3
POLY-THEGO STANDARD: PRT: 773 AA.
AC P56689;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (TO POL).
GN POL OR POLA.
OS Thermococcus gorgonarius.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID-71997;
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA MEDLINE-99199230; PubMed-10097083;
RA Huber R., Angerer B.;
RA "Crystal structure of a thermostable type B DNA polymerase from
RT Thermococcus gorgonarius."
RL Proc. Natl. Acad. Sci. U.S.A. 96:3600-3605(1999).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + (DNA)(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

```

```

DR PDB: 1TGO; 22-MAR-99.
DR InterPro: IPR002064; DNA_pol_B.
DR Pfam: PF00136; DNA_pol_B_1.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLBc; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme;
KW 3D-structure.
FT DISULFID 428 442
FT DISULFID 506 509
SQ SEQUENCE 773 AA; 89812 MW; F67AF04E875FE844 CRC64;

```

Query Match 80.8%; Score 3253.5; DB 1; Length 773;
 Best Local Similarity 78.3%; Pred No. 7.4e-176;
 Matches 605; Conservative 83; Mismatches 82; Indels 3; Gaps 2;

```

QY 1 MIFDDYITKDGKPIIRIKKENGFEKIELDPHOPYIALKDSALDEIKAKGERHG 60
Db 1 MILDFTYTEDGKPVIRIKKENGFEKIDYDRNEEYIYALLKDSALIDVKKITAEHRG 60
QY 61 KTVRVYDAVKVKKFLGRDVEVWKLFEPHPOVPAKRIIRHPAVIYDIEYDIPAKRY 120
Db 61 TVRVVRAEKVKKFLGRPIEVWKLFETHPOVPAIRDKIKHPAVVYDIEYDIPAKRY 120
QY 121 LIDKGLIPMEGDEELKMAFDIETFEHGEDEFGKEIIMISYADEBEARVITWKNDLPY 180
Db 121 LIDKGLIPMEGDEELKMAFDIETLYHEGEFEFAGSIIMISYADEGAVITWKNDLPY 180
QY 181 VDVSNEREMIKRFVOIVREKDPVLITYNGDNFDLPYLKRAEKLGVTLLGRDKEHE 240
Db 181 VDVSTEKEMIKRFVIVREKDPVLITYNGDNFDPAVLKRSSEKIGVFIIIGR-EGSE 238
QY 241 PKTHMGDSFAVEIKRIHFEDLPVVRRTINLPTYLEVYEAIVGKTSKIGAEIATL 300
Db 239 PKIQRGDFFAVEVGRHIFDLPVVRRTINLPTYLEVYEAIVGQPEKVAEELTAQ 298
QY 301 WTEESMKRLAQYSMEDARATYELKEFPMEALAKIGOSVWDVSRSTGNLYEMVLL 360
Db 299 WETGEGLEAVARYSMDADVTEELKEFPMEAGLSRLVGSIMVDSRSTGNLYEMVLL 358
QY 361 RYAVERNELAPNKPDEEYRRLRTTYLGIVYKEPRLGEMNTIYIDFCLYPSIIVHN 420
Db 359 RYAVERNELAPNKPDERELARR-RESYAGGYKEPERGLMENIVYIDFSLYPSIITHN 417
QY 421 VSPDITLEREGCKNYDAPIVGYKFCDFPGFIPSTIGELITMRKOEIKKKMAATIDPEKK 480
Db 418 VSPDITLNRGCEYDAPVGHKFCDFPGFIPSLGLDLLEEROKKKMAATIDPEKK 477
QY 481 MLYROAVAKLANSYGYMGYPKARWYSKECAESVTAMGRHYEMTLREIEKKGFLV 540
Db 478 LLDYORAIKILANSYGYGYAKARWYCKECAESVTAMGRQYIETTLREIEKKGFLV 537
QY 541 YADDTGFAVITIGEKPEYTKKAKKEFLKINSKLPGLLELEEGFYLRGFVAKKRYAVI 600
Db 538 YADDTGFAVITIGADEYTKKAKKEFLIDYINAKLPGLLELEEGFYLRGFVAKKRYAVI 597
QY 601 DEGRITRTGLVVRDMSVETKETOAKVLAIIKDSVEKAVETKDVVEELAKVOYL 660
Db 598 DEEDKTTTGLLEIVRDMSEIKETOARVLEAILKHGDEEAVRYLKEVTELSKTEVPP 657
QY 661 EKLVIHEQITKDLSEYKAGISPHVATLAKRAGIKVRPCTIISYVLRGSGKISDRVILL 720
Db 658 EKLVIHEQITRDLKOVKAGISPHVATLAKRAGIKRPTVISTYVLRGSGKISDRVILL 717
QY 721 SEYDPKKHRYDPDYIENQVLPVLRILAEAGYRKEDLKYSKQVGLDANLKK 773
Db 718 DEFDAKHRYDAEYIENQVLPVLRILAEAGYRKEDLKYSKQVGLDANLKK 770

```

RESULT 4

DR PROSITE: P500116; DNA-POLYMERASE_B.1.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 FT CONFLICT 263 263 V -> A (IN REF. 1).
 FT CONFLICT 267 267 A -> T (IN REF. 1).
 FT CONFLICT 277 277 A -> V (IN REF. 1).
 FT CONFLICT 281 281 F -> S (IN REF. 1).
 FT CONFLICT 320 320 Q -> H (IN REF. 1).
 FT CONFLICT 339 339 R -> T (IN REF. 1).
 FT CONFLICT 359 359 K -> N (IN REF. 1).
 FT CONFLICT 391 391 K -> R (IN REF. 1).
 FT CONFLICT 756 756
 SQ SEQUENCE 771 AA: 89496 MW: 1104870458A5522 CRC64:

Query Match 78.6%; Score 3164.5; DB 1; Length 771;
 Best local Similarity 76.1%; Pred. No. 7.5e-171;
 Matches 588; Conservative 89; Mismatches 93; Indels 3; Gaps 2;

QY 1 MFDVTYITKDGPIIRIFKKEGKELIDPHFOPIYALAKDSDAIDEIKAKGERHG 60
 DB 1 MIDADYITFDGPIIRIFKKEGKELIDPHFOPIYALAKDSDAIDEIKAKGERHG 60
 QY KIYRVDAVVKKKFLGRPLEVKKLLEHPDPAIRKREHPAVDIYEYDIPFAKRY 120
 DB 61 KIYRVDAVVKKKFLGRPLEVKKLLEHPDPAIRKREHPAVDIYEYDIPFAKRY 120
 QY 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 DB 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 QY 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 DB 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 QY 181 VDVVSNREMIKRPVOIVREKDPDLITVNGDNFPLYLKRAEKAGVTLGRDKEHPE 240
 DB 181 VDVVSNREMIKRPVOIVREKDPDLITVNGDNFPLYLKRAEKAGVTLGRDKEHPE 240
 DB 181 VDVVSNREMIKRPVOIVREKDPDLITVNGDNFPLYLKRAEKAGVTLGRDKEHPE 240
 QY 241 PKIHRMDSFAVEIKGRIHDLFPVVRTINLTLYLEAYEAVLTKSKLGAETIAI 300
 DB 241 PKIHRMDSFAVEIKGRIHDLFPVVRTINLTLYLEAYEAVLTKSKLGAETIAI 300
 DB 239 PKMQRGDSLAVELKIGRIHDLFPVVRTINLTLYLEAYEAVLTKSKLGAETIAI 298
 QY 301 WETEESKKLAQYSMEDARATYELGKEFFPMEALAKLIGOSVWDVSRSTGNLVEMWYL 360
 DB 299 WETGKGLERVAKYSMEDAKVTFLGKEFFPMEALAKLIGOSVWDVSRSTGNLVEMWYL 358
 QY 361 RYAVERNELAPNKPDDEEYRRRLRTYLGKYVKEPERGMENTIYDFRCLYPSIITVHN 420
 DB 359 RYAVERNELAPNKPDDEEYRRRLRTYLGKYVKEPERGMENTIYDFRCLYPSIITVHN 418
 QY 421 VSPDTLEREGCKNVDAPIYGYKFCDFPGFISILGELITMROEIKKKKATIDPIEKK 480
 DB 419 VSPDTLEREGCKNVDAPIYGYKFCDFPGFISILGELITMROEIKKKKATIDPIEKK 478
 QY 481 MLDYRORAVKLANSYGYGYGPRKAMYKSCAESVTAMGRHYIEMTIKIEEKGFKVL 540
 DB 479 MLDYRORAVKLANSYGYGYGPRKAMYKSCAESVTAMGRHYIEMTIKIEEKGFKVL 537
 QY 541 YATDDEYVATIPBEKPTIKKAKKELKYNSKLPGLLEEGEYFARGFVTKKKKAL 600
 DB 538 YATDDEYVATIPBEKPTIKKAKKELKYNSKLPGLLEEGEYFARGFVTKKKKAL 597
 QY 601 DEGRATITRLEVRVRSMSIAKETQAKVLEALIKEDSVKAEIYVADVVEELAKYVPL 660
 DB 598 DEGRATITRLEVRVRSMSIAKETQAKVLEALIKEDSVKAEIYVADVVEELAKYVPL 657
 QY 661 EKLVIHBOITKDSLEYKKAIPHVAIAKRLAKIGIKVARGTIIISYVLRSKGISDRVYL 720
 DB 658 EKLVIHBOITKDSLEYKKAIPHVAIAKRLAKIGIKVARGTIIISYVLRSKGISDRVYL 717
 QY 721 SEVDPRKKAKDPDYIYINOVLPVAVLRLLEAFGRKEDLKQSSQOVLDAMK 773
 DB 718 SEVDPRKKAKDPDYIYINOVLPVAVLRLLEAFGRKEDLKQSSQOVLDAMK 770

ID DPOL_PYRSE STANDARD: PRT: 771 AA.
 AC P77932;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL.
 OS Pyrococcus sp. (strain GE23).
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_Taxid=6036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Camdon M., Querellon J., Bouyoub A., Raguene G., Barbier G.,
 RA Forterre P., Dietrich J.;
 RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 + (DNA)(N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 254173; CAA90887.1; -
 CC HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_POL_B.
 DR Pfam: PF00136; DNA_POL_B.1.
 DR Pfam: PF03104; DNA_POL_B-exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLB; 1.
 DR PROSITE: P500116; DNA-POLYMERASE_B.1.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome;
 SQ SEQUENCE 771 AA: 89400 MW: 1506756996056C8 CRC64:

Query Match 78.6%; Score 3163.5; DB 1; Length 771;
 Best local Similarity 75.9%; Pred. No. 8.5e-171;
 Matches 587; Conservative 89; Mismatches 94; Indels 3; Gaps 2;

QY 1 MFDVTYITKDGPIIRIFKKEGKELIDPHFOPIYALAKDSDAIDEIKAKGERHG 60
 DB 1 MIDADYITFDGPIIRIFKKEGKELIDPHFOPIYALAKDSDAIDEIKAKGERHG 60
 QY KIYRVDAVVKKKFLGRPLEVKKLLEHPDPAIRKREHPAVDIYEYDIPFAKRY 120
 DB 61 KIYRVDAVVKKKFLGRPLEVKKLLEHPDPAIRKREHPAVDIYEYDIPFAKRY 120
 QY 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 DB 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 QY 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 DB 121 LIDKGLPMGDELKLMADIEFYHGEDEFKGEIIMISYADEEARVITKNDLPY 180
 QY 181 VDVVSNREMIKRPVOIVREKDPDLITVNGDNFPLYLKRAEKAGVTLGRDKEHPE 240
 DB 181 VDVVSNREMIKRPVOIVREKDPDLITVNGDNFPLYLKRAEKAGVTLGRDKEHPE 240
 DB 181 VDVVSNREMIKRPVOIVREKDPDLITVNGDNFPLYLKRAEKAGVTLGRDKEHPE 240
 QY 241 PKIHRMDSFAVEIKGRIHDLFPVVRTINLTLYLEAYEAVLTKSKLGAETIAI 300
 DB 239 PKMQRGDSLAVELKIGRIHDLFPVVRTINLTLYLEAYEAVLTKSKLGAETIAI 298
 QY 301 WETEESKKLAQYSMEDARATYELGKEFFPMEALAKLIGOSVWDVSRSTGNLVEMWYL 360
 DB 299 WETGKGLERVAKYSMEDAKVTFLGKEFFPMEALAKLIGOSVWDVSRSTGNLVEMWYL 358
 QY 361 RYAVERNELAPNKPDDEEYRRRLRTYLGKYVKEPERGMENTIYDFRCLYPSIITVHN 420
 DB 359 RYAVERNELAPNKPDDEEYRRRLRTYLGKYVKEPERGMENTIYDFRCLYPSIITVHN 418
 QY 421 VSPDTLEREGCKNVDAPIYGYKFCDFPGFISILGELITMROEIKKKKATIDPIEKK 480

```

Db 419 VSPDTLNRENCKEYDVAPOVGHRCCKDFPGFIPSLGLNLEBRÖKIKRKMESKDPVEKK 478
Oy 481 MLDYQRAVKLANSYGYGMPKARWYSKECAESVTANGRIHYIEMTTEIEKKGFKVL 540
Db 479 LLDYQRAVKLANSYGYGMPKARWYSKECAESVTANGRIHYIEMTTEIEKKGFKVL 537
Oy 541 YADTGAFATIPGEKPEFKKAKFEKLYKINSKLPGLLEEGFYLGFEFAKKRYAVI 600
Db 538 YIDTGLATITIGAKPNEKEKALKFVEITNSKLPGLLEEGFYLGFEFAKKRYAVI 597
Oy 601 DEGRITRTGLEVRDRMSSEIAKETQAKVLEALKEDESVKAVEIVKDVIEIARYOVL 660
Db 598 DEGRITRTGLEVRDRMSSEIAKETQAKVLEALKEDESVKAVEIVKDVIEIARYOVL 657
Oy 661 EKLVIHEQTKLSEYKAGPHVAIAKRLAAGIVRPGTISYLVKSGKISDRVILL 720
Db 658 EKLVIHEQTKLSEYKAGPHVAIAKRLAAGIVRPGTISYLVKSGKISDRVILL 717
Oy 721 SEYDPRKKHKYDDPYIENOVLPALVLEAFGRKEDLKYOSSKOVGLDPAWLK 773
Db 718 EEFDRKKHKYDDPYIENOVLPALVLEAFGRKEDLKYOSSKOVGLDPAWLK 770

```

```

RESULT 7
DPOL_PYRFU STANDARD: PRG: 775 AA.
AC P80061: P95584:
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Pfu polymerase).
GN POL OR PF0212.
OS Pyrococcus furiosus, and
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261, 2262;
RN 11
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC SPECIES=P. furiosus; STRAIN=VCI / DSM 3638;
RX MEDLINE=93181200; PubMed=8441634;
RA Uemori T., Ishino Y., Toh H., Asada F., Kato I.;
RT "Organization and nucleotide sequence of the DNA polymerase gene from
RL Nucleic Acids Res. 21:259-265(1993).
RN 12
RP SEQUENCE OF 396-418 FROM N.A.
RC SPECIES=P. furiosus;
RX MEDLINE=92107689; PubMed=1762925;
RA Mathur E.J., Adams M.W., Callen W.N., Cline J.M.;
RT "The DNA polymerase gene from the hyperthermophilic marine
RL archaeobacterium, Pyrococcus furiosus, shows sequence homology with
RL alpha-like DNA polymerases."
RN 13
RP SEQUENCE OF 396-502 FROM N.A.
RC SPECIES=P. furiosus;
RA Mathur E.J., Adams M.W., Callen W.N., Cline J.M.;
RL Submitted (Nov-1991) to the EMBL/Genbank/DBJ databases.
RN 14
RP SEQUENCE FROM N.A.
RC SPECIES=P. woeisel; STRAIN=DSM 3773;
RX MEDLINE=98432910; PubMed=9758761;
RA Dabrowski S., Kur J.;
RT "Cloning and expression in Escherichia coli of the recombinant his-
RL tagged DNA polymerases from Pyrococcus furiosus and Pyrococcus
RN protein Expr. Purif. 14:131-138(1998).
RN 15
RP SIMILARITY TO OTHER POLYMERASES.
RX MEDLINE=92253396; PubMed=1579479;
RA Forterre P.;
RT "The DNA polymerase from the archaeobacterium Pyrococcus furiosus does

```

```

RT not testify for a specific relationship between archaeobacteria and
RL eukaryotes."
RC Nucleic Acids Res. 20:1811-1811(1992).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- SUBUNIT: MONOMER.
CC -1- BIOTECHNOLOGY: Because Pfu DNA polymerase exhibits the lowest
CC error rate of any thermostable DNA polymerase studied, it is
CC routinely used for PCR. It is sold by Promega.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: D12983; BA002362.1; -.
DR EMBL: U84155; AB07984.1; -.
DR PIR: S35543; S35543.
DR HSRP: P56689; ITGO.
DR InterPro: IPR002064; DNA_POL_B.
DR Pfam: PF00136; DNA_POL_B_1.
DR Pfam: PF03104; DNA_POL_B_exo; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLBc; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
DR TRANSFERASE: DNA-directed DNA polymerase; DNA replication;
DR DNA-binding; Hydroxylase; Endonuclease.
DR KW
DR SEQUENCE 775 AA; 90113 MW; 5668E3D9F4DBD40E CRC64;
SQ

```

```

Query Match 78.3%; Score 3151; DB 1; Length 775;
Best Local Similarity 74.6%; Pred. NO. 4.3e-170;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;
Oy 1 MIFPDYITKDGKPIIRIFKENGKFEKLEDPHOPYIYALKDSADIEIKAIKGERHG 60
Db 1 MILDVDYITEEGKPIYRIFKENGKFEKLEHDPFRPYIYALRDSKIEEVKIKTGERRG 60
Oy 61 KIVRVDAVKYKKFELGRVVEYKLIIEHPQVPLRGRIRRHPAVIDIYEDIFPARY 120
Db 61 KIVRIYDAVKYKKFELGRVVEYKLIIEHPQVPLRGRIRRHPAVIDIYEDIFPARY 120
Oy 121 LIDKGLIPWEGDEELKMAFDIEFYEHEGDEKGEIIMISYADEEAVITWKNIDLPY 180
Db 121 LIDKGLIPWEGDEELKMAFDIEFYEHEGDEKGEIIMISYADEEAVITWKNIDLPY 180
Oy 181 VDVSNEREMIKRVOYVIREKQDVLYTNGDFDLPYIKRAEKLVTLLGRKHEPE 240
Db 181 VEVVSSEMERIKRFRIRIREKDPDIIVYNGDSFDPYLAKEKIKGLITIGRGS--E 238
Oy 241 PKIHMGDSFAVEIKGRHFDLFPVVRRTINPTYLEAVYAVIGKTKSGAEIAT 300
Db 241 PKMORTIGMTAVEKGRHFDLHYITRTINPTYLEAVYAVIGKTKSGAEIAT 298
Oy 301 WETESMKKLAQVSMEDARATYELKEPPEMEALAKLIGOSVMDVSRSSTGNLVEWYL 360
Db 299 WESGLENLVAKYSMEDAKATYELKEPPEMEALAKLIGOSVMDVSRSSTGNLVEWFL 358
Oy 361 RVAYERNELAPKPDPEEYRRRLRTTYLAGYKKEPBGKLWENITVLDRCCLPSITVHN 420
Db 359 RKAYERNELAPKPDPEEYRRRLRTTYLAGYKKEPBGKLWENITVLDRCCLPSITVHN 418
Oy 421 VSPDTLREGCNKYDVAIPVGVGKPKDFGFIPTSLGELITKROEIKKMKATDIPK 480
Db 419 VSPDTLNLEGCNKYDIAPOVGHRCCKDFPGFIPSLGLNLEBRÖKIKRKMESKDPVEKK 478
Oy 481 MLDYQRAVKLANSYGYGMPKARWYSKECAESVTANGRIHYIEMTTEIEKKGFKVL 540

```

```

Db 479 LLDYRQAKILANLSFGYGYGAKARWYCEKCAESYAMARKYIELWKKLEEFKPVLL 538
QY 541 YADTDGYYATIPGPKPTTIKKAKKFLKYINSKLPGLLEEGYRGFEVAKKRAVI 600
Db 539 YIDTDGLATIPGGESEIKKALEPKYINSKLPGLLEEGYRGFEVAKKRAVI 598
QY 601 DEGRITTRGLEVVRDMSIAKETQAKVLEALKEDSVKAEIYKDVYEIAKYQVPL 660
Db 599 DEGRVTRRGLLEIVRRDMSIAKETQAKVLEALKEDSVKAEIYKDVYEIAKYQVPL 658
QY 661 EKLVIHQIKDSEYKATIPHVAIAKRLAKIKYAPGIIISYIVLRGSGKISDRVIL 720
Db 659 EKLATIEQITRPLEHYKKAIPHVAIAKRLAKIKYAPGIIISYIVLRGSGKISDRVIL 718
QY 721 SEYDPKKHKYDPYIENQVLPVLRILEAFYGRKEDLKQSSQKQVGLDAML 772
Db 719 EYDPRKKHKYDAEYIENQVLPVLRILEAFYGRKEDLKQSSQKQVGLDAML 770

```

RESULT 8

DPOL.PYRSD STANDARD; PRT; 1312 AA.

```

AC 051334; 051335; 051336;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Deep vent DNA polymerase) [Contains:
GN Endonuclease P1-Papi (EC 3.1.-.-) (Psp-GDB pol intein)].
OS Pyrococcus sp. (strain GB-D).
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_Taxid=69013;
RN [1]
RX MEDLINE=94094330; PubMed=8269515;
RA Xu M.-Q., Southworth M.W., Mersha F.B., Hornstra L.J., Peller F.B.;
RT Identification of a branched intermediate."
RL Cell 75:1371-1377(1993).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
CC INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
CC ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
CC INTEIN.
CC -1- CATALYTIC ACTIVITY: N deoxy nucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION.
CC -1- BIOCHEMISTRY: Used in the PCR method because of its high
CC thermostability and low error rate. Sold by New England Biolabs.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U00707; AAA67130.1; -
CC EMBL: U00707; AAA67131.1; -
CC EMBL: U00707; AAA67132.1; -
CC HSSP: P56689; ITGO.
CC REBASE: 2619; PI-Papi.
CC InterPro: IPR002064; DNA_pol_B.
CC InterPro: IPR003586; HincC.
CC InterPro: IPR003587; HincN.

```

```

DR InterPro: IPR00203; Intein.
DR InterPro: IPR004042; Intein_endonuc.
DR Pfam: PF00136; DNA_pol_B_2.
DR Pfam: PF03104; DNA_pol_B_exo_1.
DR PRINTS: PR00106; DNAPOLB.
DR PRINTS: PR00379; INTEIN.
DR SMART: SM00305; HincC_1.
DR SMART: SM00306; HincN_1.
DR SMART: SM00486; POLBc_1.
DR PROSITE: PS00116; DNA_POLYMERASE_B_1.
DR PROSITE: PS00881; PROTEIN_SPLICING_1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
KW Protein splicing; Intein homing.
FT CHAIN 1 492 DNA POLYMERASE, 1ST PART.
FT CHAIN 493 1029 ENDONUCLEASE P1-PSPI.
FT CHAIN 1030 1312 DNA POLYMERASE, 2ND PART.
SQ SEQUENCE 1312 AA; 152852 MW; B62518805641D26A CRC64;

```

Query Match

Best Local Similarity 72.9%; Score 2934.5; DB 1; Length 1312;

```

Matches 597; Conservative 86; Mismatches 87; Indels 539; Gaps 3;
QY 1 MFDTYITKDGKPIIRIFKENGKELIDPHQPIYALDKDSADIEKAIKGERRG 60
Db 1 MILDADYITEDGKPIIRIFKENGKELIDPHQPIYALDKDSADIEKAIKGERRG 60
QY 61 KIVRVDAVYKKKFGROVYKKLFEHPQDVPALRGKREHVAIDIVYDIPAKRY 120
Db 61 KIVRIIDAEKVRKKKFGRLPEVWKLFHEHPQDVPALRGKREHVAIDIVYDIPAKRY 120
QY 121 LIDKGLIPMGDEELKLMFADIEFYHGEDEFGKELIMISYADEEARVITKKNIDLY 180
Db 121 LIDKGLIPMGDEELKLMFADIEFYHGEDEFGKELIMISYADEEARVITKKNIDLY 180
QY 181 VDVVSNREBKIKFVQIVRKPDVLYITNGDNFDLPYLKRAKAGVILLGRDKEHE 240
Db 181 VDVVSNREBKIKFVQIVRKPDVLYITNGDNFDLPYLKRAKAGVILLGRDKEHE 240
QY 181 VEYVSSEREMIKRFLVIRKEDPDVITYNGSFDLPYLKRAKAGVILLGRDKEHE 238
Db 181 VEYVSSEREMIKRFLVIRKEDPDVITYNGSFDLPYLKRAKAGVILLGRDKEHE 238
QY 241 PKIHRGDSFAVYIKIRIHFDLPVVRRTINLPYTLVAVYEVNLTAKSKLCAEETIAI 300
Db 241 PKIHRGDSFAVYIKIRIHFDLPVVRRTINLPYTLVAVYEVNLTAKSKLCAEETIAI 300
QY 239 PKMQRGDMVAIKIRIHFDLVHVRITNLPYTLVAVYEVNLTAKSKLCAEETIAI 298
Db 239 PKMQRGDMVAIKIRIHFDLVHVRITNLPYTLVAVYEVNLTAKSKLCAEETIAI 298
QY 301 WETESMKLIAQYSMEDARATYELGKEFPEMEALAKLIGQSVADVSRSSTGNLVEMYL 360
Db 301 WETESMKLIAQYSMEDARATYELGKEFPEMEALAKLIGQSVADVSRSSTGNLVEMYL 360
QY 299 WETGKGLERVAKYSMEDAKYTELGREFFPEMAQLSRVGOPLMDVSRSTGNLVEMYL 358
Db 299 WETGKGLERVAKYSMEDAKYTELGREFFPEMAQLSRVGOPLMDVSRSTGNLVEMYL 358
QY 361 RVAYERNELAPNKPDEEYRRLRTYVLGYVKEPERGMENTYIDFCLYPSITVTEN 420
Db 361 RVAYERNELAPNKPDEEYRRLRTYVLGYVKEPERGMENTYIDFCLYPSITVTEN 420
QY 359 RKAYERNELAPNKPDEREYRRLRESYAGYVKEPERGMENTYIDFCLYPSITVTEN 418
Db 359 RKAYERNELAPNKPDEREYRRLRESYAGYVKEPERGMENTYIDFCLYPSITVTEN 418
QY 421 VSPDTERECCKNYDAVPIGYKCFDPGIFISIGELITMROETIKKMKATIDPEKK 480
Db 421 VSPDTERECCKNYDAVPIGYKCFDPGIFISIGELITMROETIKKMKATIDPEKK 480
QY 419 VSPDTERECCKNYDAVPIGYKCFDPGIFISIGELITMROETIKKMKATIDPEKK 478
Db 419 VSPDTERECCKNYDAVPIGYKCFDPGIFISIGELITMROETIKKMKATIDPEKK 478
QY 481 LMDYKQRAVYK----- 490
Db 481 LMDYKQRAVYK----- 490
QY 479 LMDYRQAKILANLSILPEEWVPLIKNGKVIKIRIGDFVGLAKANQGVKKTGTDEVLE 538
Db 479 LMDYRQAKILANLSILPEEWVPLIKNGKVIKIRIGDFVGLAKANQGVKKTGTDEVLE 538
QY 491 ---LHA----- 493
Db 491 ---LHA----- 493
QY 539 VAGIHAFSEDRKSKKARVAVKAVIRHRYSGNVYRIVLANSGRKITTEGHSLEFVRNGDL 598
Db 539 VAGIHAFSEDRKSKKARVAVKAVIRHRYSGNVYRIVLANSGRKITTEGHSLEFVRNGDL 598
QY 494 ----- 493
Db 494 ----- 493
QY 599 VEATGEDVKIGDGLAVPRSVNLPKERNLIVLLNLSPEETEDIIITPYVGRKNFK 658
Db 599 VEATGEDVKIGDGLAVPRSVNLPKERNLIVLLNLSPEETEDIIITPYVGRKNFK 658
QY 494 ----- 493
Db 494 ----- 493
QY 659 GMLRTLRLWIFGEKRVATASRYLRHLNGLYIRLRKIGVDIIDKEGLEKRYTLYEKLVYD 718
Db 659 GMLRTLRLWIFGEKRVATASRYLRHLNGLYIRLRKIGVDIIDKEGLEKRYTLYEKLVYD 718
QY 494 ----- 493
Db 494 ----- 493

```

```

Db 719 VRYNGKREYIWEFNAVDVLSIMPEELKEWRIGTRNGFMGTVEIDEDFAKLLGYV 778
Qy 494 ----- 493
Db 779 SEGSAKKKNOTGMSYVRLYNENDEVLDMHLLAKKFGCKVRKKNVETIPKMAVYI 838
Qy 494 ----- 493
Db 839 FESLOGLTAENKRVPEVIFTSSKGVNAAFLEGYFIGDDGVHPSKRVRLSTSELYNGLV 898
Qy 494 ----- 493
Db 899 LLNLSGVSAIKLGYDSGYRVYVNEELKTEYRKKKNVSHIVPKDILKETFGKVFQK 958
Qy 494 ----- 493
Db 959 NISYKKEPRLVENGKLDREKAKRIEMLLNGDIVDRAVEIKREYYDGYVDLSVDEENF 1018
Qy 494 ----- 493
Db 1019 LAGFGLVAHNHYGYGYAKARWCKCAESVTAMGREYIEFYAKKLEEKGEFVLYID 1078
Qy 544 TDGFYATIPGEKPEPTIKKAKAEFLYINSKLPGLLELEYEGFYLRGFFVAKKRVAVIDEE 603
Db 1079 TDGLYATIPGAKPEEIKKALEFVDYINAKLPGLELEYEGFYLRGFFVAKKRVAVIDEE 1138
Qy 604 GRITRGLGVYRRDSEIATAKETOAKVLEAILKEDSVKAVETIKVDVNEELAKYOVLEK 663
Db 1139 GKITRGLGVYRRDSEIATAKETOAKVLEAILKEDSVKAVETIKVDVNEELAKYOVLEK 1198
Qy 664 VIHEOITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIYSYIVLNGSKISDRVILLSEY 723
Db 1199 VIYEGITRPLHEHYKAIQPHVAIAKRLAKGIVRPGTIIYSYIVLNGSKISDRVILLSEY 1258
Qy 724 DPKKKYOPDYIENOVPAVIRLIEAFGYRREDLKYYSSKQVGDAML 772
Db 1259 DLRRKHDAEYIENOVPAVIRLIEAFGYRREDLKYYSSKQVGDAML 1307

RESULT 9
DPOL_PRRHO STANDARD: PRT: 1235 AA.
AC 059610:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [contains: Pho pol intein (Pho Pol I
  intein)].
GN POL OR PH1947 OR PHBTO47.
OS Pyrococcus horikoshii.
OC Archaea: Euryarchaeota: Thermococcales: Thermococcaceae: Pyrococcus.
OX NCBI_TaxId=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
  Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
  Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohnuku Y.,
  Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
  Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
  Masuchi Y., Shizuya H., Kikuchi H.;
RA "Complete sequence and gene organization of the genome of a hyper-
  thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RT DNA Res. 5:55-76(1998).
RL
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
  + [DNA](N).
CC -I- PM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
  A POST-TRANSLATIONAL EXCISION OF THE INTERFERING REGION (INTEIN)
  FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000007; BAA1074.1; -
DR HSSP: P56689; INGO.
DR InterPro: IPR002064; DNA_POL_B.
DR InterPro: IPR003586; HincC.
DR InterPro: IPR003587; HincN.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR004042; Intein_endonuc.
DR Pfam: PF00136; DNA_POL_B; 2.
DR Pfam: PF0104; DNA_POL_B_exo; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00305; HincC; 1.
DR SMART: SM00306; HincN; 1.
DR SMART: SM00486; POLB; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
DR PROSITE: PS00881; PROTEIN_SPLICING; 1.
DR TRANSFERASE: DNA-directed DNA polymerase: DNA replication;
KW DNA-binding; Autocatalytic cleavage; Protein splicing;
FT CHAIN 1 492 DNA POLYMERASE, 1ST PART (POTENTIAL).
FT CHAIN 493 952 PHO POL INTEIN (POTENTIAL).
FT CHAIN 953 1235 DNA POLYMERASE, 2ND PART (POTENTIAL).
SO SEQUENCE 1235 AA; 143086 MW; 73CC7AA14873CCE4 CRC64;

```

```

Query Match 71.7%; Score 2886; DB 1; Length 1235;
Best Local Similarity 46.5%; Pred. No. 6,5e-155;
Matches 574; Conservative 107; Mismatches 89; Indels 464; Gaps 5;

Qy 1 MIFDYITTKDGKPIIRIFKKEGSEFKIELDPHQPIYVALLKDDSAIDEIKAKGERHG 60
Db 1 MILDADYITTEODGPIIRIFKKEGSEFKVEYDNRFPYIALLRDDSADIDEIKAKIQAHRG 60

Qy 61 KIVRVYDAVKKKKKGLGDEVYWKLFEPHPQVPAALRKIRHPAVIDIEYEDIPFARY 120
Db 61 KVRIVETEKIQRKPLGRPIEYWKLYLHPQVPAIRKIRHPAVVIDIEFEDIPFARY 120

Qy 121 LIDKGLIMEGDEELKIAFDIEFYHGGDEFKGEIIMISVADDEEAVITMKNDIPY 180
Db 121 LIDKGLITMEGNEKLTPLAVDIEITYHGGDEGKGPVIMISVADDEGAKVITMKNDIPY 180

Qy 181 DVVYSNEREMIKRFVQIVREKDPDLITYNGDNFDPYLIRKAEKLGVTLLGRDKHEP 240
Db 181 VEVVYSEREMIKRLIRVIRKEKDPDVIITYNGDNFDPYLIRKAEKLGVTLLGRD -NSE 238

Qy 241 PKIRMGSPAVEIKGRHFDLPFVVRTINLPYTLAVYEAUVGKRSKGAETIAI 300
Db 241 PKMOKGDSLAVEIKGRHFDLPFVIRITINLPYTLAVYEAIVGKREKRYADEIKA 298

Qy 301 WETESMKLLQVSMEDARATYELGKEFFPMAEIAKLIGSVMPVSSSTGNLEWYLL 360
Db 301 WETGSGLEIRVAKYSMEDAKVITELGREFFPMQAOLARLVGPVMPVSSSTGNLEWYLL 358

Qy 361 RYAVERNELAPKPDDEEYRRRLRTTYLGYYKEPREGMENVITLDFCLYPSITVTHN 420
Db 361 RYAVERNELAPKPDDEEYRRRLRTTYLGYYKEPREGMENVITLDFCLYPSITVTHN 418

Qy 421 VSPDLEEGCKNDYVAIVGKFKKDPGFPISLIGELITMRQIKKMKATIDPIEK 480
Db 421 VSPDLEEGCEYDAVAVGKHFCKDPGFPISLIGOLEEROKIKRMKESKDPVEK 478

Qy 481 MIDYRQRAVKL----- 491
Db 479 LLDYRQRAIKLIANSILPDEMLPIYENKRVKIGDIDREIENARVARDGETELE 538
Qy 492 ----- 491

```



```

Db 539 VKDLKALSFNRETKSEIKKVKYKALIRHRYSGKYVSIKLSGRIKITSGHLSFSVKNKGL 598
QY 492 ----- 491
Db 599 VKVAGDELKPGDLVYVPGRLKLPESKQVNLNVELLKLPEETSNIYMIPIVKGRKNFEK 658
QY 492 ----- 491
Db 659 GMLKTLVIFGEGERPRTAGRYLKHLERLGYVKLKRRCCEVLDMESLKRKYKLEYLKN 718
QY 492 ----- 491
Db 719 LKYNAGSRATWVEFNSLRDYSMLPIELKEMITGEPRPKIGTFIDVDSFAKLGYI 778
QY 492 ----- 493
Db 779 SSGDVEKDRVKFHSKDDQNVLEDIAKLAKLFGKYRGRGYIEVSGKISHAIFRYLAGSKR 838
QY 494 ----- 493
Db 839 IPEFIPTSPMDIKVAFKLGNGNAEBLTFSTKSELVNLILLINSIGVSDIKIEHEKV 898
QY 494 ----- NSYYGYM 500
Db 899 YRYVYINKKSSNGDIVDSVESIEVEKEYGYVDLSVEDNENFLVGFLLYAHNSYGY 958
QY 501 GYPAKAWYSKCAESVTAMGRHYIEMTKEIEKFGKVLADTDGFYATIGEKRP-ETI 559
Db 959 GYAKARWYKCEKCAESVTAMGRHYIEMTKEIEKFGKVLADTDGFYATIGEKRP-ETI 559
QY 560 KKKAKELKATINSKLPBLLLEVEGFLRGFPYAKKRYAVIDEGRITTTGEGLEVYRDM 619
Db 1018 KRRLEFVDYINSKLPBLLLEVEGFLRGFPYAKKRYAVIDEGRITTTGEGLEVYRDM 1077
QY 620 EIAKETQAKYLEALIKEDSYEKAVEIYKDVVEIAKQVPLEKVIHEOTKDLSEKAI 679
Db 1078 EIAKETQAKYLEALIKEDSYEKAVEIYKDVVEIAKQVPLEKVIHEOTKDLSEKAI 679
QY 680 GPHVAIAKRLANGIKVPRPTIISYIVLRGSKISDRVILLSEYDPKHKRYDPDYIENQ 739
Db 1138 GPHVAIAKRLANGIKVPRPTIISYIVLRGSKISDRVILLSEYDPKHKRYDPDYIENQ 739
QY 740 VLPVAVILILEAFGRKEDLTKYSSKQVGLDAMIK 773
Db 1198 VLPVAVILILEAFGRKEDLTKYSSKQVGLDAMIK 773
RESULT 10
POLYMERASE
OC Archaea: Thermococcus; Thermococcales; Thermococcaceae; Thermococcus.
OC NCBI:TaxID=46540;
RN SEQUENCE FROM N.A.
RP STRAIN-ST557;
RA Camdon M., Querellou J.;
RL Submitted (Feb-1996) to the EMBL/Genbank/DBJ databases.
RN CHARACTERIZATION OF INTEINS.
RP STRAIN-ST557;
RA MEDLINE=20112788; PubMed=10644683;
RA Saves I., Ozanne V., Dietrich J., Masson J.-M.;
RT "Inteins of Thermococcus funticolans DNA polymerase are endonucleases

```

```

RT with distinct enzymatic behaviors."
RL J. Biol. Chem. 275:2335-2341(2000).
CC -1- FUNCTION: PI-TfuI recognizes and cleaves a minimal sequence of 16
CC base pairs (bp) on supercoiled DNA with either Mn(2+) or Mg(2+) as
CC cofactor. It cleaves linear DNA only with Mn(2+) and requires a
CC 19-bp minimal recognition sequence. The optimal temperature for
CC activity is 70 degrees Celsius.
CC -1- FUNCTION: PI-TfuII is a highly active homing endonuclease using
CC Mg(2+) as cofactor. Its minimal recognition and cleavage site is
CC 21 bp long either on linear or circular DNA substrates. Its
CC endonuclease activity is strongly inhibited by the 3' digestion
CC product, which remains bound to the enzyme after the cleavage
CC reaction. The optimal temperature for activity is 70 degrees
CC Celsius.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION
CC (INTEINS) FOLLOWED BY PEPTIDE LIGATION.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -1- SIMILARITY: IN THE INTEIN SECTION, BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC
CC EMBL: Z69882; CAA3738.1; -
CC DR HSSP: P56689; ITGO.
CC DR REBASE: 4500; PI-TfuI.
CC DR REBASE: 4501; PI-TfuII.
CC DR InterPro: IPR002064; DNA_POL_B.
CC DR InterPro: IPR003586; HINTC.
CC DR InterPro: IPR003587; HINTN.
CC DR InterPro: IPR002203; INTEIN.
CC DR InterPro: IPR004042; INTEIN_endonuc.
CC DR Pfam: PF00136; DNA_POL_B_3.
CC DR Pfam: PF03104; DNA_POL_B_exo_1.
CC DR PRINTS: PR00379; INTEIN.
CC DR SMART: SM00305; HINTC; 2.
CC DR SMART: SM00306; HINTN; 1.
CC DR SMART: SM00486; POLBC; 1.
CC DR PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
CC DR PROSITE: PS00881; PROTEIN_SPLICING; 2.
CC KW transferase: DNA-directed DNA polymerase; DNA replication;
CC DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
CC protein splicing; Intron homing; Magnesium; Manganese.
CC FT CHAIN 1 406
CC FT CHAIN 1 766
CC FT CHAIN 1 900
CC FT CHAIN 901 1282
CC FT CHAIN 1283 1523
CC FT CHAIN 1523 175917 MW: 7A2AC8236B2E5F5 CRC64;
CC SQ SEQUENCE 1523 AA; 175917 MW; 7A2AC8236B2E5F5 CRC64;
Query Match 69.8%; Score 2809; DB 1; Length 1523;
Best Local Similarity 39.0%; Pred. No. 1,8e-150;
Matches 593; Conservative 86; Mismatches 91; Indels 752; Gaps 4;
QY 1 MIFDDYITKDGKPIRIKKEGKFEIIEPHFQYIYALLKDSADIEIKAKGERHG 60
DB 1 MILDPTYITEDGPRVIRVKKKEGKFEIIEPHFQYIYALLKDSADIEIKAKGERHG 60
QY 61 KIVRVDAVKKKKGLGRVYWKLLFEHPQDVPAALRGKIRHPAVIDYEXDIPAKRY 120
DB 61 TTVRVVRAKGVKKKGLGRVYWKLLFEHPQDVPAALRGKIRHPAVIDYEXDIPAKRY 120
QY 121 LIDKGLIPGEGELKMAFDIETFYHEDGEGKGIIMISVADDEAVITWKNIDLPY 180
DB 121 LIDKGLIPGEGELKMAFDIETFYHEDGEGKGIIMISVADDEAVITWKNIDLPY 180

```


D	b	121	LIDGILPMDEBELKMLADIDETLVHEGEEFEGEILMIISVADEBGARVITWKKIDLP	180
O	y	181	VDVVSNEREMIKRFVQIVIREKDPVLITYNQDNFDLPYLKRAEKLGVTLLGRDKEHE	240
D	b	181	VDVVSFEREMIKRFKLKVKKEKDPVLITYNQDNFDPAVLKRSSEKLGVFILGRGS--E	238
O	y	241	PKIHRMDSFAVELKGIHEDLPVVRRTNLPTYLLEAVYEAVLGKTKSKGAELAI	300
D	b	239	PKIQRMDRFAVEYKGIHFDLPVIRHTINLPTYLLEAVYEAIFGQPEKYEAEIQA	298
O	y	301	WETESMKKLAQYSMEDARATYELGKEFFPEMAELKLGQSVMPYSRSTGNLEWYLL	360
D	b	299	WETEGLEIARARYSMEAKATYELGKEFFPEMAQSLRLVGQSFMDYSRSTGNLEWYLL	358
O	y	361	RVAERNELAPNKPDDEEYERRLTTYLGQYKEPERGIMENITYLDFRC-----	410
D	b	359	RKAERNELAPNKPSGRELERR-RGGYAGGYVEPERGIMENIAYLDFRCHADKTVLK	417
O	y	411	-----	410
D	b	418	GKGVANISEVREGDYVLGIDGQKVORVWEYDEGELVINGLKCTPNHKLPVVRTERQ	477
O	y	411	-----	410
D	b	478	TAIRDSLAKSFLTKKVKGKLITTPLEFKIGKITEREDVPEEBILKGBLAGIILABETLLRK	537
O	y	411	-----	410
D	b	538	DVEYFDSRGKRRVSHQYRVEITVGAQEDFORIYIYIPERLEGYTPSYRRKKNATFE	597
O	y	411	-----	410
D	b	598	KVAKEYYLRVREIMDGIENLHAPSVLRGFEEGDSVNRKRTVYVNOGTNNEMKIEVVS	657
O	y	411	-----	410
D	b	658	KLLKLGIPHRRTTYDTEKEKMTHTHILEIAGROSLIFQTTVGFISTEKMALEEAIR	717
O	y	411	-----LYPSITVTNN	420
D	b	718	NREVRLENNAFYTLADFTAKTEYKGYKYDILTLEGTPTYPPANGILTHNSIXPSIISHN	777
O	y	421	VSPOTLEBEGCKATDAPRYGVAFCDPFGFIPSLILGELITRKQELKKKKATIDPLEKK	480
D	b	778	VSPOTLREBEGGEYDAPQVGRHCFCDPFGFIPSLILGDLDEROKVKKMKATVDPLEKK	837
O	y	481	MLDYROBAVKLHANSYGVYGVYKAPKARYSECAESVTAMGRHYIEMTKEIEKEGFEVL	540
D	b	838	LLDYRORALIKLANSEFYGYGVYKAKARYCECAESVTAMGRQYIETTMKEIEKGFEVL	897
O	y	541	YAD-----	543
D	b	898	YADSVTGDTEVTIRNRGRIEFVPIEKLPERVDRHVGKEKEYVLGGVEALTLDNRRRLWK	957
O	y	544	-----	543
D	b	958	KVPYVMRHKTDKRIYRVWFTNSWYLDVTEHDSLIGYLTNSKVPKPKLERLVEYKPEEL	1017
O	y	544	-----	543
D	b	1018	GKKVKSLLTPNRPARTIKANPIAVKLMBELIGLVGDGNGWQOSMWAKYVYGLSCGIDKA	1077
O	y	544	-----	543
D	b	1078	EIERKVYLPLEASVINSYNDKSKKDVSILSKWMLAGFVWKFKDBNGNKALPSPFMFLP	1137
O	y	544	-----	543
D	b	1138	REYIEAFRLGLFSADGTVSLRGIGPEIRLTSVNBRELSAVRKLMLVGVSNSLFTETKPN	1197
O	y	544	-----	543
D	b	1198	RYLEKESGTSIHVIRLKNKHFRPADRIGFLIDRKSSTKLSJENSGHTNKKRAYKYDDEVLVP	1257

```

OY 544 -----TDGYATITPEKPEPTIKKAKKEFLKYN 571
Db 1258 RKIEETITDGYVYDIEVECTHRFPFANGILIVHNTDGEFAITPGADATVYKKAKEFLNYIN 1317
OY 572 SKLPGLLELEGEYGLRGEFVAKKRAVYIDEGRITTRGLAEVRNDSSEIAKETQAKYLE 631
Db 1318 PRLPGLLELEGEYGFRRGFFVAKKAAVYIDEGSKITTRLEIYRNRMSVAKETQARVLE 1377
OY 632 ALLKEDSVKAEIVKDVVEEIAKYOVPLEKLYVIEHQITKIDSEYKAISGPHVAIAKRLAA 691
Db 1378 ALLRGDVEAEVARIKEVTEKISKYEVPPEKLVIEHQITRELKDYKATGPHVAIAKRLAA 1437
OY 692 KRIKVRPGGIISYIVLRSGKISDRVILLSEKDPKKHKADDPXYIENOVLPVLRILEAF 751
Db 1438 KRIKVRPGGVISYIVLKSGGRIGDRITPDEPDPKHRYDAEYIENOVLPVLRILKAF 1497
OY 752 GYRKEDLKYOSKOYGLDAMLK 773
Db 1498 GYKKEDLRKYQKTRGYGLAMLK 1519

RESULT 11
DPOL_PYRO STANDARD: PRT: 1671 AA.
AC P77933:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase [EC 2.7.7.7] [contains: Endonuclease PI-PKOI
DE (EC 3.1.-.-) (Pko pol-1 intein) (IVS-A); Endonuclease PI-PKOI
DE (EC 3.1.-.-) (Pko pol-2 intein) (IVS-B)].
GN
OS Pyrococcus kodakaraensis.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN-KOD1:
RX MEDLINE=98027387; PubMed=9361436;
RA Takagi M., Nishio M., Kakihara H., Kitabayashi M., Inoue H.,
RA Kawakami B., Oka M., Imanaka T.;
RT "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1
RT and its application to PCR.";
RL Appl. Environ. Microbiol. 63:4504-4510(1997).
RN [2]
RP CHARACTERIZATION OF INTEINS.
RC STRAIN-KOD1;
RX MEDLINE=98416198; PubMed=9742242;
RA Nishio M., Fujiwara S., Takagi M., Imanaka T.;
RT "Characterization of two intein homing endonucleases encoded in the
RT DNA polymerase gene of Pyrococcus kodakaraensis strain KOD1.";
RL Nucleic Acids Res. 26:4409-4412(1998).
CC -I- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
CC INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
CC ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
CC INTEIN. PI-PKOI RECOGNIZES 5'-GATTGATCCGCTGACC-3' AND PI-PKOII
CC RECOGNIZES 5'-CAGCTACTACGGTTC-3'. BOTH ARE THERMOSTABLE.
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -I- PWM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (PONENTIAL).
CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -I- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)

```

CC or send an email to license@slb-slb.ch).

DR EMBL: D29671; BAA06142.2; .
DR HSSP: P56689; ITGO.
DR REBASE: 3792; PI-PKOI.
DR REBASE: 3793; PI-PKOI.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR003586; HincC.
DR InterPro: IPR003587; HincN.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR004042; Intein_endonuc.
DR Pfam: PF00136; DNA_pol_B_3.
DR Pfam: PF03104; DNA_pol_B_exo_1.
DR PRINTS: PR00379; Intein.
DR SMART: SM00305; HincC_2.
DR SMART: SM00306; HincN_2.
DR SMART: SM00486; POLBc_1.
DR PROSITE: PS00116; DNA_POLYMERASE_B_1.
DR PROSITE: PS00881; PROTEIN_SPLICING_2.
KW Transferrase; DNA-directed DNA polymerase; DNA replication;
M DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
N Protein splicing; Intron homing.
FT CHAIN 1 406 DNA POLYMERASE, 1ST PART.
FT CHAIN 407 766 ENDONUCLEASE PI-PKO I.
FT CHAIN 767 851 DNA POLYMERASE, 2ND PART.
FT CHAIN 852 1388 ENDONUCLEASE PI-PKO II.
FT CHAIN 1389 1671 DNA POLYMERASE, 3RD PART.
SQ SEQUENCE 1671 AA; 193490 MW; 4A17F1C8120EE455 CRC64;

Query Match 69.1%; Score 2782; DB 1; Length 1671;
Best Local Similarity 36.0%; Pred. No. 6.9e-149;
Matches 602; Conservative 82; Mismatches 86; Indels 900; Gaps 4;

QY 1 MIDDTYITKDGKPIIRIKFKENGFEKIELDPHOPITYALLKDDSAIDEIKATKGRHG 60
1 MIDDTYITKDGKPIIRIKFKENGFEKIELDPHOPITYALLKDDSAIDEIKATKGRHG 60
DB 1 MIDDTYITKDGKPIIRIKFKENGFEKIELDPHOPITYALLKDDSAIDEIKATKGRHG 60

QY 61 KIYRVDAVKKKKFLGRDVEYKKLIFEHPQDVPALEKGIKREHPAVIDIYEDIPAKRY 120
61 KIYRVDAVKKKKFLGRDVEYKKLIFEHPQDVPALEKGIKREHPAVIDIYEDIPAKRY 120
DB 61 KIYRVDAVKKKKFLGRDVEYKKLIFEHPQDVPALEKGIKREHPAVIDIYEDIPAKRY 120

QY 121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEPFKGETIMISYDEEGARITYTKNDLPY 180
121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEPFKGETIMISYDEEGARITYTKNDLPY 180
DB 121 LIDKGLIPMEGDEELKLMADIEFTYHEGDEPFKGETIMISYDEEGARITYTKNDLPY 180

QY 181 VDVVSNEREMIKRFOIVREKDPDLITYNGDNFDLPYLTKRAEKLGVTLTGRDKEHP 240
181 VDVVSNEREMIKRFOIVREKDPDLITYNGDNFDLPYLTKRAEKLGVTLTGRDKEHP 240
DB 181 VDVVSNEREMIKRFOIVREKDPDLITYNGDNFDLPYLTKRAEKLGVTLTGRDKEHP 240

QY 241 PRTIRMGDSFAVEIKRIRHFDLPVVRRTINLPYTLLEAVYEAVLKTSKLGAEETAI 300
241 PRTIRMGDSFAVEIKRIRHFDLPVVRRTINLPYTLLEAVYEAVLKTSKLGAEETAI 300
DB 241 PRTIRMGDSFAVEIKRIRHFDLPVVRRTINLPYTLLEAVYEAVLKTSKLGAEETAI 300

QY 239 PKIQRMGDRFAVEIKRIRHFDLPVVRRTINLPYTLLEAVYEAVFGQPKREKVAEETIT 298
239 PKIQRMGDRFAVEIKRIRHFDLPVVRRTINLPYTLLEAVYEAVFGQPKREKVAEETIT 298
DB 239 PKIQRMGDRFAVEIKRIRHFDLPVVRRTINLPYTLLEAVYEAVFGQPKREKVAEETIT 298

QY 301 WEFEESMKKIAQSMEDARTYELGKEFFPMEALAKLIGOSWMDVSRSTGLVEMYL 360
301 WEFEESMKKIAQSMEDARTYELGKEFFPMEALAKLIGOSWMDVSRSTGLVEMYL 360
DB 301 WEFEESMKKIAQSMEDARTYELGKEFFPMEALAKLIGOSWMDVSRSTGLVEMYL 360

QY 299 WEFEESMKKIAQSMEDARTYELGKEFFPMEALAKLIGOSWMDVSRSTGLVEMYL 358
299 WEFEESMKKIAQSMEDARTYELGKEFFPMEALAKLIGOSWMDVSRSTGLVEMYL 358
DB 299 WEFEESMKKIAQSMEDARTYELGKEFFPMEALAKLIGOSWMDVSRSTGLVEMYL 358

QY 361 RYAVENNELAPNKPDEEYRRLRTTYLGYKVEPERGLMENTIYDFFC----- 410
361 RYAVENNELAPNKPDEEYRRLRTTYLGYKVEPERGLMENTIYDFFC----- 410
DB 361 RYAVENNELAPNKPDEEYRRLRTTYLGYKVEPERGLMENTIYDFFC----- 410

QY 359 RKAYENNELAPNKPDEELARR-RQSYEGGYKVEPERGLMENTIYDFFC----- 417
359 RKAYENNELAPNKPDEELARR-RQSYEGGYKVEPERGLMENTIYDFFC----- 417
DB 359 RKAYENNELAPNKPDEELARR-RQSYEGGYKVEPERGLMENTIYDFFC----- 417

QY 411 ----- 410
411 ----- 410
DB 411 ----- 410

QY 418 GKGIIINSEQEGDYVLIDGMQVRKRWEDYKGETLVNINGLKCTPNHKLPAVTKNERG 477
418 GKGIIINSEQEGDYVLIDGMQVRKRWEDYKGETLVNINGLKCTPNHKLPAVTKNERG 477
DB 418 GKGIIINSEQEGDYVLIDGMQVRKRWEDYKGETLVNINGLKCTPNHKLPAVTKNERG 477

QY 411 ----- 410
411 ----- 410
DB 411 ----- 410

QY 478 TRIROSLANSFLTKKVKGIITTPLEYEIGRATSENIPEBEVLKGLAGIILAEGLTLRK 537
478 TRIROSLANSFLTKKVKGIITTPLEYEIGRATSENIPEBEVLKGLAGIILAEGLTLRK 537
DB 478 TRIROSLANSFLTKKVKGIITTPLEYEIGRATSENIPEBEVLKGLAGIILAEGLTLRK 537

QY 411 ----- 410
411 ----- 410
DB 411 ----- 410

QY 538 DVEYFDSKKRKRRIHQRYVEITIGKDEEFRDRTIYIFERLFGITPISSEKGTNAVTL 597
538 DVEYFDSKKRKRRIHQRYVEITIGKDEEFRDRTIYIFERLFGITPISSEKGTNAVTL 597
DB 538 DVEYFDSKKRKRRIHQRYVEITIGKDEEFRDRTIYIFERLFGITPISSEKGTNAVTL 597

QY 411 ----- 410
411 ----- 410
DB 411 ----- 410

QY 598 KVAKKNVYLKVKELMONIESLHPSVLRGFEEEDGSVNRVRSIVATQCTKNEMKIKLVS 657
598 KVAKKNVYLKVKELMONIESLHPSVLRGFEEEDGSVNRVRSIVATQCTKNEMKIKLVS 657
DB 598 KVAKKNVYLKVKELMONIESLHPSVLRGFEEEDGSVNRVRSIVATQCTKNEMKIKLVS 657

QY 411 ----- 410
411 ----- 410
DB 411 ----- 410

QY 658 KLSOLGIPHOYTYOYQENGKDRSRYILEITGKDLILFQTLGIFISERRNALLNKALS 717
658 KLSOLGIPHOYTYOYQENGKDRSRYILEITGKDLILFQTLGIFISERRNALLNKALS 717
DB 658 KLSOLGIPHOYTYOYQENGKDRSRYILEITGKDLILFQTLGIFISERRNALLNKALS 717

QY 411 ----- 410
411 ----- 410
DB 411 ----- 410

QY 718 QRENNLENNGFYRLSEFNVSTYEYCKVYDLEGTPIYFANGILLHNSLYPSIITTHN 777
718 QRENNLENNGFYRLSEFNVSTYEYCKVYDLEGTPIYFANGILLHNSLYPSIITTHN 777
DB 718 QRENNLENNGFYRLSEFNVSTYEYCKVYDLEGTPIYFANGILLHNSLYPSIITTHN 777

QY 421 VSPDTEREGCKNVDYAPVIGYKCDPFGFISILGELITMROEIKKKKATIDPIEKK 480
421 VSPDTEREGCKNVDYAPVIGYKCDPFGFISILGELITMROEIKKKKATIDPIEKK 480
DB 421 VSPDTEREGCKNVDYAPVIGYKCDPFGFISILGELITMROEIKKKKATIDPIEKK 480

QY 778 VSPDTEREGCKNVDYAPVIGYKCDPFGFISILGELITMROEIKKKKATIDPIEKK 837
778 VSPDTEREGCKNVDYAPVIGYKCDPFGFISILGELITMROEIKKKKATIDPIEKK 837
DB 778 VSPDTEREGCKNVDYAPVIGYKCDPFGFISILGELITMROEIKKKKATIDPIEKK 837

QY 481 MLDYRQRAKLAH----- 493
481 MLDYRQRAKLAH----- 493
DB 481 MLDYRQRAKLAH----- 493

QY 838 LLDYRQRAKLAH----- 493
838 LLDYRQRAKLAH----- 493
DB 838 LLDYRQRAKLAH----- 493

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 898 VSGLEVPSENRINKAKELKVKALLRHDSGKYTTIRLSGRKIKITSGHLSFVNGEL 957
898 VSGLEVPSENRINKAKELKVKALLRHDSGKYTTIRLSGRKIKITSGHLSFVNGEL 957
DB 898 VSGLEVPSENRINKAKELKVKALLRHDSGKYTTIRLSGRKIKITSGHLSFVNGEL 957

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 958 VEYTGDELKPGDLVAVPRLELPERNHVNLVELLGTPEETLDIWMIPVAKKNFK 1017
958 VEYTGDELKPGDLVAVPRLELPERNHVNLVELLGTPEETLDIWMIPVAKKNFK 1017
DB 958 VEYTGDELKPGDLVAVPRLELPERNHVNLVELLGTPEETLDIWMIPVAKKNFK 1017

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1018 GMLRTLRFEGEEKRPTARARLRLHEDLGYVRKRIGYEVLDMDSLKNRYRLYALVEN 1077
1018 GMLRTLRFEGEEKRPTARARLRLHEDLGYVRKRIGYEVLDMDSLKNRYRLYALVEN 1077
DB 1018 GMLRTLRFEGEEKRPTARARLRLHEDLGYVRKRIGYEVLDMDSLKNRYRLYALVEN 1077

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1078 VRNKGKREVLVEFNSTRDAVGLMPLEKELKEMKIGTLNGRMKLLVDESLAKLGGYV 1137
1078 VRNKGKREVLVEFNSTRDAVGLMPLEKELKEMKIGTLNGRMKLLVDESLAKLGGYV 1137
DB 1078 VRNKGKREVLVEFNSTRDAVGLMPLEKELKEMKIGTLNGRMKLLVDESLAKLGGYV 1137

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1138 SEGARKORPNKMGWSYKLYNEDPEVDDMERLASRFGKVRGRNRYEIPKIGYLL 1197
1138 SEGARKORPNKMGWSYKLYNEDPEVDDMERLASRFGKVRGRNRYEIPKIGYLL 1197
DB 1138 SEGARKORPNKMGWSYKLYNEDPEVDDMERLASRFGKVRGRNRYEIPKIGYLL 1197

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1198 FENMGVLAENKRIPEFVFTSPKGVRLALEGYFICDDGVHPKRLSTSKSELANOLY 1257
1198 FENMGVLAENKRIPEFVFTSPKGVRLALEGYFICDDGVHPKRLSTSKSELANOLY 1257
DB 1198 FENMGVLAENKRIPEFVFTSPKGVRLALEGYFICDDGVHPKRLSTSKSELANOLY 1257

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1258 LLNSYGSVAVKLGHDSGYRYVINEELPFVKDKKKNAVYSHVIRKEVLSFVGKVFQK 1317
1258 LLNSYGSVAVKLGHDSGYRYVINEELPFVKDKKKNAVYSHVIRKEVLSFVGKVFQK 1317
DB 1258 LLNSYGSVAVKLGHDSGYRYVINEELPFVKDKKKNAVYSHVIRKEVLSFVGKVFQK 1317

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1318 NVSPQFRKKNVEDGRDPEKAQRLSWLIEGDVLDLVESVDVEDYDGYVDSVEDNENF 1377
1318 NVSPQFRKKNVEDGRDPEKAQRLSWLIEGDVLDLVESVDVEDYDGYVDSVEDNENF 1377
DB 1318 NVSPQFRKKNVEDGRDPEKAQRLSWLIEGDVLDLVESVDVEDYDGYVDSVEDNENF 1377

QY 494 ----- 493
494 ----- 493
DB 494 ----- 493

QY 1378 LVGFGVLAHNSYGYGARARWYCKEASVYTAGRETIWTKIEIEKKGFKVYSD 1437
1378 LVGFGVLAHNSYGYGARARWYCKEASVYTAGRETIWTKIEIEKKGFKVYSD 1437
DB 1378 LVGFGVLAHNSYGYGARARWYCKEASVYTAGRETIWTKIEIEKKGFKVYSD 1437

QY 544 TDGFYATIPGEKPEITIKKAKELKINSKLPJLLEVEGEFVLRGFFVAKKRYAVIDE 603
544 TDGFYATIPGEKPEITIKKAKELKINSKLPJLLEVEGEFVLRGFFVAKKRYAVIDE 603
DB 544 TDGFYATIPGEKPEITIKKAKELKINSKLPJLLEVEGEFVLRGFFVAKKRYAVIDE 603

QY 1438 TDGFFATIGADDEYVKKAMEFLKYNKLPJLLEVEGEFVLRGFFVAKKRYAVIDE 1497
1438 TDGFFATIGADDEYVKKAMEFLKYNKLPJLLEVEGEFVLRGFFVAKKRYAVIDE 1497
DB 1438 TDGFFATIGADDEYVKKAMEFLKYNKLPJLLEVEGEFVLRGFFVAKKRYAVIDE 1497

QY 604 GRITRGLLEVVRDMSIAKETQAKVLEALIKEDSVKAVEIVKDYVEETIAKYOVLEK 663
604 GRITRGLLEVVRDMSIAKETQAKVLEALIKEDSVKAVEIVKDYVEETIAKYOVLEK 663
DB 604 GRITRGLLEVVRDMSIAKETQAKVLEALIKEDSVKAVEIVKDYVEETIAKYOVLEK 663

QY 1498 GKITTGLLEIVRDMSEIAKETQAKVLEALIKEDSVKAVEIVKDYVEETIAKYOVLEK 1557
1498 GKITTGLLEIVRDMSEIAKETQAKVLEALIKEDSVKAVEIVKDYVEETIAKYOVLEK 1557
DB 1498 GKITTGLLEIVRDMSEIAKETQAKVLEALIKEDSVKAVEIVKDYVEETIAKYOVLEK 1557

QY 664 VTHEQITKDLSEKAGPHVATKRLAAGIKYRPTIISYIVLRSSGKISDRVILLSEY 723
664 VTHEQITKDLSEKAGPHVATKRLAAGIKYRPTIISYIVLRSSGKISDRVILLSEY 723
DB 664 VTHEQITKDLSEKAGPHVATKRLAAGIKYRPTIISYIVLRSSGKISDRVILLSEY 723

QY 1558 VTHEQITKDLSEKAGPHVATKRLAAGIKYRPTIISYIVLRSSGKISDRVILLSEY 1617
1558 VTHEQITKDLSEKAGPHVATKRLAAGIKYRPTIISYIVLRSSGKISDRVILLSEY 1617
DB 1558 VTHEQITKDLSEKAGPHVATKRLAAGIKYRPTIISYIVLRSSGKISDRVILLSEY 1617

QY 724 DPKKHYDDPYIENOVLPVLRILLEAFGRKEDLTYOSSKOYGDAMLK 773
724 DPKKHYDDPYIENOVLPVLRILLEAFGRKEDLTYOSSKOYGDAMLK 773
DB 724 DPKKHYDDPYIENOVLPVLRILLEAFGRKEDLTYOSSKOYGDAMLK 773

QY 1618 DPKKHYDDPYIENOVLPVLRILLEAFGRKEDLTYOSSKOYGDAMLK 1667
1618 DPKKHYDDPYIENOVLPVLRILLEAFGRKEDLTYOSSKOYGDAMLK 1667
DB 1618 DPKKHYDDPYIENOVLPVLRILLEAFGRKEDLTYOSSKOYGDAMLK 1667

ID	DPOL_12	STANDARD:	PRT:	1699 AA.
DPOL_THG88	DPOL_THG88			
AC	Q9HH84:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	EC 3.1.1.- (Tsp-G68 pol-1 Intein) [Contains: Endonuclease PI-Tspg81I (EC 3.1.1.-) (Tsp-G68 pol-2 Intein)].			
DE	(EC 3.1.1.-) (Tsp-G68 pol-2 Intein)].			
GN	POL OR POL-1.			
OS	Thermococcus sp. (strain GB8).			
OC	Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.			
OX	NCBI_TaxId=105583;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Querellou J.J.E., Cambon M.A., Lessongeur F., Barbier G.;			
RT	"Thermococcales taxonomy and phylogeny based on the comparative use of			
RT	16S rDNA, 16S-23S rDNA Intergenic spacer and family B DNA polymerase			
RT	genes.";			
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE			
CC	EXHIBITS 3', TO 5' EXONUCLEASE ACTIVITY (BY SIMILARITY).			
CC	-1- FUNCTION: PI-TSPG81 AND PI-TSPG81I ARE ENDONUCLEASES			
CC	(POTENTIAL).			
CC	-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate			
CC	+ [DNA] (N).			
CC	-1- PNM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES			
CC	A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)			
CC	FOLLOWED BY PEPTIDE LIGATION.			
CC	-1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.			
CC	-1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOWING			
CC	ENDONUCLEASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: AJ250333; CAC12850.1; -			
DR	InterPro: IPR002064; DNA_POL_B.			
DR	InterPro: IPR003586; HincC.			
DR	InterPro: IPR003587; HincN.			
DR	InterPro: IPR002203; Intein.			
DR	InterPro: IPR004042; Intein_endonuc.			
DR	Pfam: PF00136; DNA_POL_B; 3.			
DR	Pfam: PF03104; DNA_POL_B_exo; 1.			
DR	PRINTS: PR00106; DNA_POLB.			
DR	PRINTS: PR00379; INTEIN.			
DR	SMART: SM00305; HincC; 2.			
DR	SMART: SM00306; HincN; 2.			
DR	SMART: SM00486; POLB; 1.			
DR	PROSITE: PS00881; PROTEIN_SPLICING; 2.			
DR	PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.			
KW	Transferase; DNA-directed DNA polymerase; DNA replication;			
KW	DNA-binding; Hydrolyase; Nuclease; Exonuclease; Multifunctional enzyme;			
KW	Protein splicing; Autocatalytic cleavage; Endonuclease; Intron homing.			
FT	CHAIN 1			
FT	CHAIN 492 1026			
FT	CHAIN 1027 1075			
FT	CHAIN 1076 1464			
FT	CHAIN 1465 1699			
SQ	SEQUENCE 1699 AA; 197323 MW; F389BA451F0B12D3 CRC64;			

Oy	1	MIFPDIYTKKGRGIIIEFKKNGEPEIIEIDPHQOYIALLKDSADIEIATIGERG	60
Db	1	MILDTDITTEGKAVIYFKKNGEKIEIYDRNPEPFLAKDSALTEBVKIKIARRG	60
Oy	61	KIVERVDAVYKKKKFIEGRDVEVWKLIJEHPQDVPALRGIRIHEHPAVIDYEIDIPFANKY	120
Db	61	TIVAVKRAEKVKKKFELGRPIEVMKLYFTTHQDVPALRODIRHEHPAVIDIYEIDIPFARKY	120
Oy	121	LIDKGLIPMGDEBLKLMAFDIEFTYHGEDSECKGEIIMISTADEEARVITWKNIIDLPY	180
Db	121	LIDKGLIPMGDEBLKLMAFDIEFTYHGESEFPAECPILMISYADEGAGAVITWKRVDPY	180
Oy	181	VDVYNSREEMKKRRVOYLVEKRDVDVITTYNGDFDLPYLITKRAEKIYUILLDGRCKEPE	240
Db	181	VDVYSTEKEMKKRFLRVYKKEKRDVDVITTYNGDFDPAYLTKRSEKGVAFILGRGS--E	238
Oy	241	PKIRMGDSFAVEIKGRIHFDLPVVARTINLPYTLAEAVYEAUIGLTKSKLGAEBIAI	300
Db	239	PKIQMGDRFAVEYKGIHFDLPVIRITINLPYTLAEAVYEAIRCKPREKAYAEIATA	298
Oy	301	WETEESKKIATQYSMEDARATYELGKEFFPEMAELAKIGQSYMDVSNSTGNLVWYLL	360
Db	299	WETEGELEARYASMEBAKVTPELGKEFFPEMAQLSRLIGQSLMDVSNSTGNLVWFL	358
Oy	361	RVAERNEMLANPKRDEEYERRLTYTIGGYVKEPERGMENTIYIDFCLYPSIITVHN	420
Db	359	RKAEERNEMLANPKRDERELAR--ROSYAGGYVEPERGLMNNIVYLDPSILPSIITHN	417
Oy	421	VSPPTLEBEGCKNDVAPIYGYKFKCKPFPISIGELITMROEIKKKAKATIPIEKK	480
Db	418	VSPPTLRECKEKEDVAPQYGHKFCCKDPFIPSLIGDILBEROKIKKKRATIDPEKK	477
Oy	481	MLDYROBRAVLAH-----	493
Db	478	LIDYROBAIKILANSILPDEMLPLVNGRLKLYRIGDVPDNTMKKGQPLENDGTEVLEYS	537
Oy	494	-----	493
Db	538	GIEAISERNKTKIAELIKPVAKALIRHRYRCKVYDIKLSGRNIKYTEGHSLEAFRDELVE	597
Oy	494	-----	493
Db	598	VTGGEIKRGPDIANPRRVNLPENHERINLIEILLGPPESTDYULTIPYKRRKKFPFGM	657
Oy	494	-----	493
Db	658	LRTLRFIEEBQRPRTARYLEHLQKGYVKLKRAYEIVNKREALRNYRKKLEYLAEVRK	717
Oy	494	-----	493
Db	718	YNGKREYLVHFNDLRNEIKFMPDEBLEBKVGTLNGFMEBPIIEVGEDFAKLLGYVSE	777
Oy	494	-----	493
Db	778	GYARKQNRQNGMSYSVKIYNNDQRYLDMEKILASKFPGVRBGKNYVEISKKMAVLE	837
Oy	494	-----	493
Db	838	SICGTLAENKRVPEVITSPESVRMAFEBEYFIGDDDLHPSKRVRLSTSEELVNGVYL	897
Oy	494	-----	493
Db	898	INSIGISAIRIPDSGVYRVLVNDELPFLGNRRKKRKNAAYSHVLPKEIIEETPGOKFQKM	957
Oy	494	-----	493
Db	958	SPAFLNKEVEKGEIDACKARIRIAMLLEBGDIVDRVEKVTVEDYEGYVDLSVEENENFLA	1017
Oy	494	-----	543
Db	1018	GFGMLYAHNSYGYGYAKARAWYCRGCAESVTAMGRSYIETITRIELEKFGKVLVYADSV	1077
Oy	544	-----	543

Db 1078 AGNTEIIRNGKVEFVEPIEKLFRVDYRIGKEKECALGVEALLTDNRGLRWKRPYI 1137
 QY 544 ----- 543
 Db 1138 MRHKTKKRYWFTNSWYTDVTDHSLIGYLTNSKYKSEKPLERLVEYKPRELGEKVK 1197
 QY 544 ----- 543
 Db 1198 SLITLRAIARSIAKANDIAVRLMELLGLVGDGNGGSHKMAKYVGLSCGLDKAEIEEK 1257
 QY 544 ----- 543
 Db 1258 VLRLPLKAGIISNYGSKKGVDSISKWLAFWVKYFKDENGKRIKIPSEFNLPREYIE 1317
 QY 544 ----- 543
 Db 1318 AFLRLGFSADGTVSLRGIIEIRLTSVRLSNEVRKLLMLVGSNSMFTETTPNKYLG 1377
 QY 544 ----- 543
 Db 1378 ESGTRSIHVRIKNNHFRFAKRIEFLDRKATKLSDNLRHTNKKMAYDEFLVYPKIEE 1437
 QY 544 ----- 543
 Db 1438 INTDRYVYDIEVEGTRHFFANGILVNTDGFATIPGADAEYVKKAMEFLKTIYNAKLP 1497
 QY 577 LLELEVEGYLRFYFAKRRKAAVYIDEEGRITTTGLEVVRDSEIAKETAQVLEALKE 636
 Db 1498 LLELEVEGYVRFVYFAKRRKAAVYIDEEGRITTTGLEVVRDSEIAKETAQVLEALKE 1557
 QY 637 DSVKAEVYKDVVEELAKVQVLEKVIHEQITKLSKRAIGPHVIAKRLAAKIKV 696
 Db 1558 GDVEEAIRYIKVEYETKSKYEPKELVHEQITRDLKDYKAPHAVAKRLAAKIKI 1617
 QY 697 RPTGIISYIYLRGSGKISDVILLSEYDPKKHXYDPDYIENOVLPVILILEAFGRKE 756
 Db 1618 RPTGIISYIYLRGSGKISDVILLSEYDPKKHXYDPDYIENOVLPVILILEAFGRKE 1677
 QY 757 DLRYQSSQYVGLDAMLK 773
 Db 1678 DLRYQSSQYVGLDAMLK 1694
 RESULT 13
 DPOL_THERHY STANDARD: PRT; 1668 AA.
 AC 09H05;
 RT 16-OCT-2001 (Rel. 40, Last Created)
 RT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) [Contains: Endonuclease PI-TyII
 (EC 3.1.-.-) (Thy pol-1 intein); Endonuclease PI-TyI (EC 3.1.-.-)
 (Thy pol-2 intein)] (Fragment).
 GN POL.
 OS Thermococcus hydrothermalis.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
 OX NCBI_TaxID=46539;
 RN (1) SEQUENCE FROM N.A.
 RA Querrelou J.J.E., Camhon M.A., Leongeur F.O., Barbier G.;
 RT "Thermococcales taxonomy and phylogeny based on the comparative use of
 RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
 RT genes.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN RP CHARACTERIZATION OF PI-TYII.
 RX MEDLINE=20512590; PubMed=11058140;
 RA Saves I., Eleaume H., Dietrich J., Masson J.-M.;
 RT "The Thy pol-2 intein of Thermococcus hydrothermalis is an
 RT isochloromer of PI-TyI and PI-TyII endonucleases.";
 RL Nucleic Acids Res. 28:4391-4396(2000)
 -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE

CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY (BY SIMILARITY).
 CC -1- FUNCTION: PI-TYI AND PI-TYII ARE ENDONUCLEASES. PI-TYI CLEAVES
 CC THE INTERNESS SEQUENCE OF THE THY DNA POL GENE. IT REQUIRES A 21-
 CC BP MINIMAL RECOGNITION SEQUENCE.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + (DNA)(N).
 CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
 CC ENDONUCLEASE FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AJ245819; CAC18555.1; -.
 CC REBASE: 4832; PI-TyI.
 DR InterPro: IPR002064; DNA_POL_B.
 DR InterPro: IPR003586; HINIC.
 DR InterPro: IPR003587; HININ.
 DR InterPro: IPR002203; InteIn.
 DR InterPro: IPR004042; InteIn_endonuc.
 DR Pfam: PF00136; DNA_POL_B_3.
 DR Pfam: PF03104; DNA_POL_B_exo_1.
 DR PRINTS: PRO0106; DNAPOLB.
 DR PRINTS: PRO0379; INTEIN.
 DR SMART: SM00305; HINIC_2.
 DR SMART: SM00306; HINIC_2.
 DR SMART: SM00486; POLBC_1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
 DR PROSITE: PS00881; PROTEIN_SPLICING_1.
 DR TRANSFERASE: DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Hydrolyase; Nuclease; Exonuclease; Multifunctional enzyme;
 KW Protein splicing; Autocatalytic cleavage; Endonuclease; Intron homing.
 FT NON_TER 1 1
 FT CHAIN 1 458 DNA POLYMERASE, 1ST PART.
 FT FT 459 995 ENDONUCLEASE PI-TYI (POTENTIAL).
 FT CHAIN 996 1044 DNA POLYMERASE, 2ND PART.
 FT CHAIN 1045 1433 ENDONUCLEASE PI-TYII (POTENTIAL).
 FT CHAIN 1434 1668 DNA POLYMERASE, 3RD PART.
 FT SEQUENCE 1668 AA; 193319 MW; 5BE805EPEDA71C8 CRC64;

Query Match 63.0%; Score 2535.5; DB 1; Length 1668;
 Best Local Similarity 33.6%; Pred. No. 5e-135;
 Matches 560; Conservative 85; Mismatches 92; Indels 929; Gaps 4;

QY 34 FQRYTALLKDSADIEIKAKGERHGIYVVDAYVKKKFLGRDVEYWKLFEEPROV 93
 Db 1 FEPYITALLKDSADIEIKAKGERHGIYVVDAYVKKKFLGRDVEYWKLFEEPROV 60
 QY 94 PALRGKIREHPAVIDIYEDIPFAKRYLDKGLIPMGDEELKMAFDIETFEHDEG 153
 Db 61 PAIRDEIRRSAAVVDIYEDIPFAKRYLDKGLIPMGDEELKMAFDIETFEHDEG 120
 QY 154 KGEIIMISVADDEEARITWKINDLPYVDVYSNEREKIRFOVIVREKDPVLTNGDN 213
 Db 121 TGPILMISVADDEEARITWKINDLPYVEVSTEKEIKRFLVREKDPVLTNGDN 180
 QY 214 FDLPLIKRAEKGVTLLGROKHEPRTKIRMGDSFAVIRGIRHEDLPVVRRTINP 273
 Db 181 FDRPVLKRCCKTGKIFTLRRDGS--EPKIQRGDPPFAVEVKRIHFDLPVVRRTINP 238
 QY 274 TYLLEAVEVAVLTKTSKGAEEIAAIWETEESMKKLAQYSMDADATYELGKEPPMDA 333
 Db 239 TYLLEAVEVAVLTKTSKGAEEIAAIWETEESMKKLAQYSMDADATYELGKEPPMDA 298


```

QY 50 EIKAIKBERG-----KIVRVDAVKVKKFL-----GRDV-----EWKLI 86
DB 60 DLKLEDEFENNTSLKVQODLTNIEIEKITYSDYIINGKDISVDFKNNKERRICKY 119
QY 87 FEHQDVPALRGKIREHPAVIDIYEDIPFAKRYLIDKGLIPM-----EGDEELK 136
DB 120 VKYNNHAKITIREYKEG---KSEFIDIPPLRRMIDODIVPSAKKYSEDKNKIDNSTEELN 176
QY 137 LMAEDIEFHEGDEFGKEGELIMISYADEBARVITKNIDLP---YVDVSNEREMIK 192
DB 177 CIAFDMELCYCKEENAKKIDIMVNLFSKIDQOKVITYKKFENSBNYCGVDVNDKEKELIQ 236
QY 193 RFQVIVAKENKDPVLITNGNFDLPYLKRAKAGVLLGLCKEHEP---KIHNRGDS 249
DB 237 KTEIL--KQYDVIYTYNGNFDLPYKKNRANIYEIDFDMASNSQOPOITIKSKGIN 294
QY 250 FAVEIKRIFHDLFPVVRRTINLPTYLLEAVYEAVLIGTKSKLGAEEIAIMETEESMKK 309
DB 295 RKSITPILIHDLPIARKLLNLTLYKLENNVOELFKINKEAVDYGDPKMKMETEDT--T 352
QY 310 LAQYMEDARATYELGKEFFMEAEALAKLIGQSVYDSSSTGNLVEMYLRLVAYERNEL 369
DB 353 LIRAYEDALYTYKMGVYFLPLEIMFSRIYNOPLDYTSRMNSQOMFELLKRSFEQNM 412
QY 370 APNKPDEEYRRRLRTYLGQYKPERGCMENITTYDFCLPYSITVHNVSPTL--E 427
DB 413 SPNPPSSSYRERAKFSYEGGYNEPLKIQEDIVSLDFMSLYPILLSHNISPEYIYE 472
QY 428 REGCKNDVAPIYGVKFCCKDFPFIPISELITMRQEKKKMKATIDPIEK-----K 480
DB 473 EKEKENNEL-----GIPIKTNELLSRKHIMLK---DKIQKNFDEEYS 516
QY 481 MLDYRQAVKLANSYGYGYPKARWYKSCAESVTAMGRHYIEMTKEIEKFGKRYL 540
DB 517 RLEHEOKSIVYLANSHGYLAFPARWYSDCAEMVTGLRKTYOETIEKAE--FGKVI 575
QY 541 YADPDGYATIPGKEPETIKK-----AKELKAYNSKILPGL 578
DB 576 YADPDGYAAMWDYKLGKCKEENDSKLSNLPKLSKEELIITTKFGLINSELPEGM 635
QY 579 ELEVEGYLGFVFAKKRYAVIDEGRITTRGLEVVRDMSLAKETQAVKLEALIKEDS 638
DB 636 ELEVEGFGKGLFVTKKRYALLIEDDGHIVKGLEVARRDSNLAKDQOAVIRALLDGD 695
QY 639 VEKAVEIVKQVVERIAKYQVPLEKVLVHEQITVDLSYKAIQPHVAIAKRLAKGIVRP 698
DB 696 VNLAKKIKITKTIDNLKGNIDKNDLHITQLTNKIEEYKSTAPHIEVAKKIKORGDSVR 755
QY 699 GTIISYVLGSGKISRVLLSEYDPKHKHYPDYIEMOVLPAVRLIEAGYKREDEL 758
DB 756 GDVISTYIVKGSRSISERAEEL--EY---ACDYDINYIIDNOVLPPVIRIMESLISSEDEL 811
QY 759 KYOSQOVGLDAML 772
DB 812 K-NSGKQFKLDQFM 824

```

```

RESULT 15
DPOL_ARCFU
ID DPOL_ARCFU STANDARD: PRT: 781 AA.
AC 029753:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN POL OR POLB AF0497.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=VC-16 / DSM 4304 / ATCC 49558;

```

```

RX MEDLINE=98049343; PubMed=9389475;
RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.R., Badger J.H., Glodex A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + (DNA)(N).
CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001070; AAB90741.1; -.
DR HSSP: P56689; ITGO.
DR TRGR: AF0497; -.
DR InterPro: IPR002064; DNA_pol_B.
DR Pfam: PF00136; DNA_pol_B; 1.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLBc; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Complete proteome.
SQ SEQUENCE 781 AA; 89850 MW; 761C7B18FCC618B8 CRC64;

```

```

Query Match 32.3%; Score 1301; DB 1; Length 781;
Best Local Similarity 37.4%; Pred. No. 5,2e-66;
Matches 305; Conservative 155; Mismatches 266; Indels 90; Gaps 21;
QY 2 IPFDYITKCGKPIIRFKENGFEKTELDPHFOPTLYALAKDSADIEKAIKGERHGK 61
DB 8 LIDADYETIGKAVVRIMCKDDGIFAYADYNDPYYIVGDE--DILKNATSTRRE 64
QY 62 IYRVYDAVKKKFLGDRVWKLIFEHPODVPALRGKIREHPAVIDIYEDIPFAKRYL 121
DB 65 VIRKSEFEKAQLTGLGREVGYIVYAHHPQHVPLKRLDYLSQFC---DVRADPIPFAVRYL 121
QY 122 IDKGL-----IPMEGDE-----ELKIMAPDIEFFHEG--DEFG 153
DB 122 IDKDLACMDGIALGEQGVINSYKIEKVERIPRMEFFPLKMLVPCCEMLSSFGPEPE 181
QY 154 KGEIIMISYADEBARVITKNIDLPYDVVSNEREMIKRFQVIVREKDPVLITNGDN 213
DB 182 KDPITIVSVKTNDDDEI-----LTGDEKTIISFVLLIKSYDDPDIIVGNDA 230
QY 214 FDLPIYLKRAEKLGVLLGLCKEHEPRTIHRMGDSFAVEIKRIFHDLFPVVRRTINLP 273
DB 231 FDMPPYLKRAERNIPLDVGRGDSN--VYFRGG---RPRITGRNLVLDIADIMRISDIK 284
QY 274 TYTLEAVYEAVLGTK- SKLGAEIAIWTESMKKLAQYMEDARATYELCKEFPFME 332
DB 285 IKRLNVAEFLGYKIELADIEADQIIRYWSRGER--EKVLAYARQDQINLYLAKELPIM 343
QY 333 AELAKLIGQSVMPDSSSTGNLVEMYLRLVAYARNELAPKPKDEEYRRRLRTYLGQYV 392
DB 344 YELSKMIRLPVDYTRMGKQYDWMLLSEAKKIGELIAPRPHAE-----SYGARY 396

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:17:02 ; Search time 23.89 Seconds
(without alignments)
313.150 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026
Sequence: 1 MIPMDYITKDGKPIIRIFK.....KEDLKYSKQVGLDAMLKK 774

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-71:**
1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	3356	83.4	1702 2	S42459 DNA-directed DNA P
2	3204.5	79.6	775 2	S67920 DNA-directed DNA P
3	3164.5	78.6	771 2	C75023 DNA polymerase I P
4	3151	78.3	775 2	S33243 DNA-directed DNA P
5	2934.5	72.9	1312 2	S68593 DNA-directed DNA P
6	2886	71.7	1235 2	C71210 Probable DNA-direc
7	2764.5	68.7	1670 2	S71551 DNA-directed DNA P
8	1301	32.3	781 2	A69312 DNA polymerase B1
9	1291	32.1	1634 2	E64410 DNA-directed DNA P
10	1209.5	30.0	586 2	C69028 DNA-dependent DNA
11	1204.5	29.9	781 2	CJ7382 DNA-directed DNA P
12	1190	29.6	803 2	B56277 DNA-directed DNA
13	1138	28.3	784 2	E72515 Probable DNA-direc
14	693.5	17.2	1088 2	T05731 DNA-directed DNA P
15	676.5	16.8	764 2	S75407 Probable DNA-direc
16	659.5	16.4	1038 1	JC5757 DNA-directed DNA P
17	657.5	16.3	1038 2	T18222 DNA polymerase del
18	657.5	16.3	1086 2	T40242 DNA polymerase del
19	649.5	16.1	1086 2	T43266 DNA-directed DNA P
20	626	15.5	901 2	E84210 DNA polymerase B1
21	621.5	15.4	1097 1	RNBRL3 DNA-directed DNA P
22	619.5	15.4	882 2	S23019 DNA-directed DNA P
23	619.5	15.4	1084 1	S19661 DNA-directed DNA P
24	616.5	15.3	872 2	JC7380 DNA-directed DNA P
25	616.5	15.3	882 2	F90201 DNA polymerase I (
26	616.5	15.3	1105 1	S40243 DNA-directed DNA P
27	614.5	15.3	1094 2	S22573 DNA-directed DNA P
28	611	15.2	1107 2	A41618 DNA-directed DNA P
29	609	15.1	1513 2	T28158 Probable DNA-direc

30	605.5	15.0	1106 1	A39299 DNA-directed DNA P
31	587	14.6	1492 2	T18560 DNA-directed DNA P
32	584.5	14.5	879 2	A56277 DNA-directed DNA P
33	583.5	14.5	1462 1	DJH0AC DNA-directed DNA P
34	581	14.4	1081 2	T20698 hypothetical prote
35	578.5	14.4	875 2	JC5186 DNA-directed DNA P
36	570.5	14.2	959 2	F72763 Probable DNA-direc
37	570	14.2	1465 2	S45628 DNA-directed DNA P
38	559.5	13.9	223 2	E69125 DNA-dependent DNA
39	548.5	13.6	1339 1	S20052 DNA-directed DNA P
40	527.5	13.1	1468 2	S58250 DNA-directed DNA P
41	522.5	13.0	1505 2	S28079 DNA-directed DNA P
42	516.5	12.8	844 2	T31321 DNA-directed DNA P
43	516	12.8	1405 1	DJ2PA DNA-directed DNA P
44	504.5	12.5	2154 2	A84669 hypothetical prote
45	500.5	12.4	913 2	T17675 DNA-directed DNA P

ALIGNMENTS

RESULT 1

S42459
DNA-directed DNA polymerase (EC 2.7.7.7) Vent, intein containing precursor - Thermoco
N:Contains: DNA endonuclease (EC 3.1.-.-) PI-I; DNA endonuclease (EC 3.1.-.-) PI-II;
C:Species: Thermococcus litoralis
C:Date: 31-Dec-1993 #sequence-revision 02-Aug-1994 #text-change 18-Jun-1999
C:Accession: S42459; S42451; S42450; S42458
R:Perler, F.B.; Comb, D.G.; Jack, W.E.; Moran, L.S.; Qiang, B.; Kucera, R.B.; Benner,
submitted to the EMBL Data Library, September 1992
A:Reference number: S42458
A:Accession: S42459
A:Molecule type: DNA
A:Residues: 1-1702 <PERI>
A:Cross-references: EMBL:M74196; NID:9154685; PID:NAA72900.1; PID:9154686
R:Perler, F.B.; Comb, D.G.; Jack, W.E.; Moran, L.S.; Qiang, B.; Kucera, R.B.; Benner,
Proc. Natl. Acad. Sci. U.S.A. 89, 5577-5581, 1992
A:Title: Intervening sequences in an Archaea DNA polymerase gene.
A:Reference number: S42450; MUID:92302285
A:Accession: S42451
A:Molecule type: DNA
A:Residues: 181-922,387-425,452-476,483-524,1021-1062,1076-1099,1466-1489,1533-1547 <
A:Cross-references: EMBL:M74198
R:Hodges, R.A.; Perler, F.B.; Noren, C.J.; Jack, W.E.
Nucleic Acids Res. 20, 6153-6157, 1992
A:Title: Protein splicing removes intervening sequences in an archaea DNA polymerase.
A:Reference number: S40788; MUID:93117082
A:Accession: S40788
A:Contents: annotation
C:Function: <VENT>
A:Description: nucleotidyltransferase
A:Note: DNA-directed DNA polymerase Vent
C:Function: <END1>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TII
C:Function: <END2>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END3>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END4>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END5>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END6>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END7>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END8>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END9>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END10>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END11>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END12>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END13>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END14>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END15>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END16>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END17>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END18>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END19>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END20>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END21>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END22>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END23>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END24>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END25>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END26>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END27>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END28>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END29>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END30>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END31>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END32>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END33>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END34>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END35>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END36>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END37>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END38>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END39>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END40>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END41>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END42>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END43>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END44>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END45>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END46>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END47>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END48>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END49>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END50>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END51>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END52>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END53>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END54>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END55>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END56>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END57>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END58>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END59>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END60>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END61>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END62>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END63>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END64>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END65>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END66>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END67>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END68>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END69>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END70>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END71>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END72>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END73>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END74>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END75>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END76>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END77>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END78>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END79>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END80>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END81>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END82>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END83>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END84>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END85>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END86>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END87>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END88>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END89>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END90>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END91>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END92>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END93>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END94>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END95>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END96>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END97>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END98>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END99>
A:Description: endonuclease; hydrolase
A:Note: DNA endonuclease PI-TIII
C:Function: <END100>

Query Match 83.4% Score 3356; DB 2; Length 1702;
Best Local Similarity 42.7% Pred. NO. 1.2e-14;
Matches 726; Conservative 33; Mismatches 15; Indels 928; Gaps 2;

QY	1	MEFDYDTIKOQKPIIRIRFFKKNSEGEFKIELDPHPQRYTALLKQDSADIELKAIKGEHNG	60
Db	1	MLDLPDITTKDOKPIIRIRFFKKNSEGEFKIELDPHPQRYTALLKQDSADIELKAIKGEHNG	60
QY	61	KIVRVAVDAVKVKKFPLGADVEMWKLIFENHPDVPALRGKIREHPAVIDIYEYDIPFAKRY	120
Db	61	KIVRVADAVKVKKFLGREVEMWKLIFENHPDVPAMRGKIREHPAVIDIYEYDIPFAKRY	120
QY	121	LIDKGLIPMEGDEELKLMAFDIETFFHEGDEFGKEITIMISYADEEARVITMKNDLPY	180
Db	121	LIDKGLIPMEGDEELKLMAFDIETFFHEGDEFGKEIIMISYADEEARVITMKNDLPY	180
QY	181	VDVVSNEREMIKRFQOIRREKODPVLITYNDNDNDPLIRAKELGVTLLGDKCHEPE	240
Db	181	VDVVSNEREMIKRFQOIRREKODPVLITYNDNDNDPLIRAKELGVTLLGDKCHEPE	240
QY	241	PKIHRMGDSFAVEIKGRHIFDLFPVVRRTINLPTYLLEAVVEAVLGKYSKLGAEETIAI	300
Db	241	PKIHRMGDSFAVEIKGRHIFDLFPVVRRTINLPTYLLEAVVEAVLGKYSKLGAEETIAI	300
QY	301	WETBESMKKLAOYSMEDARATYELKEFFPMEABELAKLIGOSVMDVSRSSTGNLVEWYL	360
Db	301	WETBESMKKLAOYSMEDARATYELKEFFPMEABELAKLIGOSVMDVSRSSTGNLVEWYL	360
QY	361	RAVYERNELAPRKPEEEYRRRLRTTYGIVGVKPEBEGMLNITYLDRCLYPSIITYHN	420
Db	361	RAVYERNELAPRKPEEEYRRRLRTTYGIVGVKPEBEGMLNITYLDRCLYPSIITYHN	420
QY	421	VSPDTLERBGCKNDYAVIYGVKPCDKDPGFIPISILGELITMRQIKKKMATDPIYEKK	480
Db	421	VSPDTLERBGCKNDYAVIYGVKPCDKDPGFIPISILGELITMRQIKKKMATDPIYEKK	480
QY	481	MLDYGQRAVKLHA-----	493
Db	481	MLDYGQRAVKLHA-----	493
QY	494	-----	499
Db	494	-----	499
QY	541	VNNLEAFSEFNKKIKESYVKKYYAKALIRHKYKKKAYEIOUSSGRKINTIAGHSLFTVRNGEI	600
Db	541	VNNLEAFSEFNKKIKESYVKKYYAKALIRHKYKKKAYEIOUSSGRKINTIAGHSLFTVRNGEI	600
QY	601	KEVSGDGIKEGDLIAPKIKILNKEGVSINIPILISDISEETADIWYTISAKRKNPFK	660
Db	601	KEVSGDGIKEGDLIAPKIKILNKEGVSINIPILISDISEETADIWYTISAKRKNPFK	660
QY	661	GMLRTLRRMFMGBENRRIRTFNKRFLYHLEKLGILKILPRGYEVTDMERLKXKXQLYERLAG	720
Db	661	GMLRTLRRMFMGBENRRIRTFNKRFLYHLEKLGILKILPRGYEVTDMERLKXKXQLYERLAG	720
QY	721	SVKTYGNKREYLWFMNETKDTISFPQKLEEMKIGTLNGFRNTCLKVDEDFSKLLGYT	780
Db	721	SVKTYGNKREYLWFMNETKDTISFPQKLEEMKIGTLNGFRNTCLKVDEDFSKLLGYT	780
QY	781	VSEGVAGAOKNKTGSIYSVLYNDDPVLBSMKNVAEKKFGXRVDRNCISIKKMAYL	840
Db	781	VSEGVAGAOKNKTGSIYSVLYNDDPVLBSMKNVAEKKFGXRVDRNCISIKKMAYL	840
QY	841	WMKCLGALAENKRIPSYILTSPBPVRWSPLEAFTGDDGDIHPSKRFRLSTKSELLANOL	900
Db	841	WMKCLGALAENKRIPSYILTSPBPVRWSPLEAFTGDDGDIHPSKRFRLSTKSELLANOL	900
QY	901	VELLSLGSISVKGIFGDSGVRYVYINEDLPQPSRSREKNYYSNLIRKELLIDYFGKEFQ	960
Db	901	VELLSLGSISVKGIFGDSGVRYVYINEDLPQPSRSREKNYYSNLIRKELLIDYFGKEFQ	960
QY	961	KNMTFKRKRELVDLSCKLMREKAKLLEPFINDDIYLDKRYKSKKEDEYCYIYDLSVENEN	1020
Db	961	KNMTFKRKRELVDLSCKLMREKAKLLEPFINDDIYLDKRYKSKKEDEYCYIYDLSVENEN	1020
QY	1021	FLVGVGLIYAHNSYGYMGYFKARWYSKECAESYTAMGRHYTEMTIHEIEKKGFVKLYA	1080
Db	1021	FLVGVGLIYAHNSYGYMGYFKARWYSKECAESYTAMGRHYTEMTIHEIEKKGFVKLYA	1080
QY	543	D-----	543

```

Db      1081 DSVSGSEIIITQNGKIRPFVKIKDLFSKVDYSIGEKYICLEGVEALITLDDGKLWMPV 1140
OY      544 ----- 543
Db      1141 PYVMRHRANKRMFRILWLTNSWTIDYTEHDSILGILNTSKTKAKKIGERLKEVKKPELGK 1200
OY      544 ----- 543
Db      1201 AVKSLICPNAPLKDENTKTSIAVKFWELVGLIVGDGMWGDSRNAEYVLGSLTGKDAEE 1260
OY      544 ----- 543
Db      1261 IKOKLLEPLKTTGVISNTPYKNEKGFNLAKSLYKFMKRHKDEKGRKRIPEFMELPV 1320
OY      544 ----- 543
Db      1321 TYIEAFRLGLFSADGVTITRKGVPEIRLNIADADFLREVRLKLIVGTSINSIFAETTPNR 1380
OY      544 ----- 543
Db      1381 YNGSVTGYSKHLRIKNKWRFAERIGFLIERKOKRLEHLKSAVYKRNITDFGFDLVHYK 1440
OY      544 ----- 543
Db      1441 KVEELIPESYVVDIEVEEETHRFPANNILVHNTDGGFYATIPGKPELIIKKAKEFLNTINS 1500
OY      573 KLPGLELELEGFYLRGFEVAKKRYAVIDEGRIITTRGLEVVRDMSIAKETQAKVLEA 632
Db      1501 KLPGLELELEGFYLRGFEVTKKRYAVIDEGRIITTRGLEVVRDMSIAKETQAKVLEA 1560
OY      633 ILEKDSVEKAAVIVADVBEIAKVOYPLEKLYIHQITKDLSEYKAIGPHVAIAKRLAK 692
Db      1561 ILEKDSVEKAAVIVADVBEIAKRYPLEKLYIHQITRDLKKAIGPHVAIAKRLAR 1620
OY      693 GKVRPGTITSYIVRGSGKISDRVILLSEYDPPKHKKDPDYIENOVLPAYLRILEAFG 752
Db      1621 GKVRPGTITSYIVRGSGKISDRVILLSEYDPPKHKKIDPDYIENOVLPAYLRILEAFG 1680
OY      753 YRKEDLKYSQKQVGLDAMLKK 774
Db      1681 YRKEDLRYSQKQVGLDAMLKK 1702

RESULT 2
S67920
DNA-directed DNA polymerase (EC 2.7.7.7) - Thermococcus sp.
C.Species: Thermococcus sp.
C.Date: 17-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 18-Jun-1999
C.Accession: S67920
R:Southworth, M.S.; Kong, H.; Kucera, R.B.; Ware, J.; Jannasch, H.W.; Perler, F.B.
submitted to the EMBL Data Library, January 1996
A.Description: Cloning, expression and modulation of the 3'-5' exonuclease activity
A.Reference number: S67920
A.Accession: S67920
A.Molecule type: DNA
A.Residues: 1-775 <SOU>
A.Cross-references: GB:047108; NID:g1197451; PIDN:AAA86769.1; PID:g1197452
A.Experimental source: strain 90N-7
A.Family: herpesvirus DNA-directed DNA polymerase
C.Keywords: DNA binding; nucleotidyltransferase

Query Match      79.6%; Score 3204.5; DB 2; Length 775;
Best Local Similarity 76.6%; Pred. No. 2.5e-157;
Matches 592; Conservative 89; Mismatches 89; Indels 3; Gaps 2;

OY      1 MFTDVTYTKDQPIIRIFKKENGSEFKIEDPHFQPIYALILKDDSAIDEIKAKGERHG 60
      ||| |||||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MILDVYITENCKPVIIRVFKKENGSEFKIEYDRTFEPYALILKDDSAIEDVKKVTAHRG 60
      ||| |||||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

OY      61 KIVRVVDAVKKKKKFLGRDVEVWKLIEHPDQVPLRLCKIKENHPAVIDITYEDIPFAKRY 120
      ||| |||||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

Db 61 TVVVKRAEKVOKKFLGRPIEWMKLYNHPDVAIRDRIRAHNAVVDIYEYDIPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMADIEIETFEHGEDEFGCEIIMISYADEEARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGDEELKLMADIEIETFEHGEDEFGCEIIMISYADEEARVITWKNIDLPY 180
QY 181 VDVAVSNEREMIKRREVOJIREKDPVLTITNGDNFDLPYLIRAEKLGVTLLGRDKEHPE 240
Db 181 VDVAVSNEREMIKRREVOJIREKDPVLTITNGDNFDLPYLIRAEKLGVTLLGRDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 300
Db 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 300
QY 239 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 298
Db 239 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 298
QY 301 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 360
Db 301 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 360
QY 299 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 358
Db 299 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 358
QY 361 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 420
Db 361 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 420
QY 359 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 417
Db 359 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 417
QY 421 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 480
Db 421 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 480
QY 418 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 477
Db 418 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 477
QY 481 MLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 540
Db 481 MLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 540
QY 478 LLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 537
Db 478 LLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 537
QY 541 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 600
Db 541 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 600
QY 538 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 597
Db 538 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 597
QY 601 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 660
Db 601 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 660
QY 598 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 657
Db 598 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 657
QY 661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 720
Db 661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 720
QY 658 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 717
Db 658 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 717
QY 721 SEYDPKKHKYDPPYIENQVLPVLRILEAFGYRKEDIKYSSKOVGLDAMLK 773
Db 721 SEYDPKKHKYDPPYIENQVLPVLRILEAFGYRKEDIKYSSKOVGLDAMLK 773

```

RESULT 3

```

C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75023
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: C75023
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-771 <RAW>
A:Cross-references: GB:AJ248288; GB:AL096836; NID:95458960; PIDN:CAB50625.1; PID:9545913
A:Experimental source: strain Orsay
C:Genetics:
C:Gene: pol1
C:Superfamily: herpesvirus DNA-directed DNA polymerase

```

```

Query Match 78.6%; Score 3164.5; DB 2; Length 771;
Best Local Similarity 76.1%; Pred. No. 2.8e-155;
Matches 588; Conservative 89; Mismatches 93; Indels 3; Gaps 2;
QY 1 MIFDPTITTKDGKPIIRIFKKEGCEKIEIDPHFOPYIYALLKDSADIEIKAIKERHG 60

```

```

Db 1 MIFDPTITTKDGKPIIRIFKKEGCEKIEIDPHFOPYIYALLKDSADIEIKAIKERHG 60
QY 61 KIVRVADVAVKAKKFLGRDVAWVKLIFEHPODVPAIIRKIRHNAVVIDIYEYDIPAKRY 120
Db 61 KIVRVADVAVKAKKFLGRDVAWVKLIFEHPODVPAIIRKIRHNAVVIDIYEYDIPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMADIEIETFEHGEDEFGCEIIMISYADEEARVITWKNIDLPY 180
Db 121 LIDKGLIPMEGDEELKLMADIEIETFEHGEDEFGCEIIMISYADEEARVITWKNIDLPY 180
QY 181 VDVAVSNEREMIKRREVOJIREKDPVLTITNGDNFDLPYLIRAEKLGVTLLGRDKEHPE 240
Db 181 VDVAVSNEREMIKRREVOJIREKDPVLTITNGDNFDLPYLIRAEKLGVTLLGRDKEHPE 240
QY 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 300
Db 241 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 300
QY 239 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 298
Db 239 PKIHRMGDSFAVEIKGRHFDLPFVVRTINLPYTTLEAVYEAVLGKTSKLGAEETAI 298
QY 301 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 360
Db 301 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 360
QY 299 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 358
Db 299 WETESMKRLAQSMEBARATYELGKEFPMEAEIATKIGOSVMDVSRSSGTGNLYEWYLL 358
QY 361 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 420
Db 361 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 420
QY 359 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 418
Db 359 RYAVERNELAPNKPDEBEYRRRLRTTYLGYYKVEBERGLMENTITLDFRCLYPSITITHN 418
QY 421 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 480
Db 421 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 480
QY 419 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 478
Db 419 VSPDTLEREGCKNTDVAPIVGYKCKDPFGFIPSLGELITMRQEIKKKKATIDPLEKK 478
QY 481 MLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 540
Db 481 MLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 540
QY 479 LLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 537
Db 479 LLDYRORAVKLANHNSYGYMGYPKARWYKSCAEVSATAMGNHYEMTKEIEKGFYVL 537
QY 541 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 600
Db 541 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 600
QY 538 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 597
Db 538 YADTDFGYATIPGEPETIKKAKAEFLKYINSKLPGLLEEGFYLRGFFVAKKRYAVI 597
QY 601 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 660
Db 601 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 660
QY 598 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 657
Db 598 DEGRITRGLLEVRRMSSEIAKETOAKVLEAILKEDSVEKAVIYVDVBEIAKYQVPL 657
QY 661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 720
Db 661 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 720
QY 658 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 717
Db 658 EKLVIHQITKDLSEYKAIQPHVAIAKRLAKGIVRPGTIIISYIVLRGSGKISDRVILL 717
QY 721 SEYDPKKHKYDPPYIENQVLPVLRILEAFGYRKEDIKYSSKOVGLDAMLK 773
Db 721 SEYDPKKHKYDPPYIENQVLPVLRILEAFGYRKEDIKYSSKOVGLDAMLK 773

```

RESULT 4

```

C:Species: Pyrococcus furiosus
C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 20-Jun-2000
C:Accession: S3543
R:Umori, T., Ishino, Y., Toh, H., Asada, K., Kato, I.
Nucleic Acids Res. 21, 259-265, 1993
A:Title: Organization and nucleotide sequence of the DNA polymerase gene from the arc
A:Reference number: S3543; MUID:93181200
A:Accession: S3543
A:Molecule type: DNA
A:Residues: 1-775 <UEM1>
A:Cross-references: EMBL:ID2983; NID:9216917; PIDN:BAA02362.1; PID:9216918
A:Accession: S44596
A:Molecule type: protein
A:Residues: 1-12 <UEM2>
C:Genetics:
C:Gene: pol
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; nucleotidyltransferase

```


Db 599 VEATGEDVKGIGDLIAVPSVNLPEKRELRNIVELLNLSPEETEDIIITIPVKGKRNFEK 658
QY 494 ----- 493
Db 659 GMLKTLRMIFGEKEKRVFASRYLRLHLENLGYIRLRKIGYDIIIDKEGLEKTYLKLVDV 718
QY 494 ----- 493
Db 719 VRVNGNREVLEFENAVRDIYSIMPEBELKEMRIGTRNGFMGTFFVIDEDFAKLGLGYV 778
QY 494 ----- 493
Db 779 SEGSAKMKKNOTGSMSTYVRLYNENDEVLDDMEHLAKKFGKVRGKNVVEIPKMAVYI 838
QY 494 ----- 493
Db 839 FESICGLTAEKNRVEVIFTSSKGVNPALEGYITGDGDVHPSKRVRLSTKSELLVNGLV 898
Y 494 ----- 493
Db 899 LLLNSLGSAILKIGDGVYRVYVNEELKFTYRKKKNVSHIVPKDILKETGKVPQK 958
QY 494 ----- 493
Db 959 NISYKKEFELVENGKLDREKAKRIEMLNGDIVLDRVVEIKREYDGYVLDVSEDENEF 1018
QY 494 ----- 493
Db 1019 LAGEFGLYAHNSYGYGYAARWYCKEASVTAMGSEYFVEYRKELEEFGRVLID 1078
QY 544 TDGEVATIPGEPETIKKAKKEFLKYNLSKPLGLELEEGFYLRGFEVAKRYAVIDEE 603
Db 1079 TDGLYATIPGAKPEIKKALEFVDYINAKPLGLELEEGFYLRGFEVAKRYAVIDEE 1138
QY 604 GRITRLEVRVRRWSEIAETQAKVLEAILKESVEKAVYKDVVEIAKYOVPLEK 663
Db 1139 GRITRLEVRVRRWSEIAETQAKVLEAILKESVEKAVYKDVVEIAKYOVPLEK 1198
QY 664 VHEIOIKRDISEYKAIQPHVAIAKRAKGIKVRGTIYIVLRSKISDRVILLSEY 723
Db 1199 VHEIOIKRDISEYKAIQPHVAIAKRAKGIKVRGTIYIVLRSKISDRVILLSEY 1258
QY 724 DPKKHXYDPDYIENOVLPVLRILAEFGYRKEDLKYOSKQVGLDAWL 772
Db 1259 DPKKHXYDPDYIENOVLPVLRILAEFGYRKEDLKYOSKQVGLDAWL 1307
RESULT 6
C71210
probable DNA-directed DNA polymerase - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: C71210
R:Kawababayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohtsuka, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137
A:Accession: C71210
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1235 <KAM>
A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA31074.1; PID:g3258391
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1947
C:Superfamily: hypothetical protein PH0202

Query Match 71.7%; Score 2886; DB 2: Length 1235;
Best Local Similarity 46.5%; Pred. No. 1,1e-140;
Matches 574; Conservative 107; Mismatches 89; Indels 464; Gaps 5;

QY 1 MIPDDYITKDKPPIIRIFKKENGKEFEKIELDPHPQVYIALLKDDSAIDELKAIKGERRG 60
Db 1 MILDADYITTEGKPIIRIFKKENGKEFEKYEDNRPYIYALLRDSADIDELKAIKAIQNHG 60
QY 61 KIVRVDAVAVKVKKFLGDVEMVKLIFEHPQDVPAIRGKIREHPAVIDIYEXDIPFARKY 120
Db 61 KIVRVDAVAVKVKKFLGDVEMVKLIFEHPQDVPAIRGKIREHPAVIDIYEXDIPFARKY 120
QY 121 LIDKGLIPMEDEELKMAPIETFFYHGEDEFGKEIIMISYADEEKARVITWKIDLPY 180
Db 121 LIDKGLIPMEDEELKMAPIETFFYHGEDEFGKEIIMISYADEEKARVITWKIDLPY 180
QY 181 VDVVSNREMIKRFVQVIRKEDPDVLYNDGNFPLIRIKRAEKLGTLLGRDKEPE 240
Db 181 VDVVSNREMIKRFVQVIRKEDPDVLYNDGNFPLIRIKRAEKLGTLLGRDKEPE 240
QY 241 PKIRHMGDSFAVEIKGRIFHFDLPVVRRTINLPYTLAEVYEAVALGKTRSKGAEIAAI 300
Db 241 PKIRHMGDSFAVEIKGRIFHFDLPVVRRTINLPYTLAEVYEAVALGKTRSKGAEIAAI 300
QY 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEIAKLIGOSWQDVSSSGNLYEWLL 360
Db 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEIAKLIGOSWQDVSSSGNLYEWLL 360
QY 361 RYAVERNELAPNKPEDEEYRRRLRTTYLGYYKPEBEGMLNITYLDPRCLYPSIYTHN 420
Db 361 RYAVERNELAPNKPEDEEYRRRLRTTYLGYYKPEBEGMLNITYLDPRCLYPSIYTHN 420
QY 421 VSPDLTREGCKNDVADIVGYKFCDEPGFIPSLIGELITWROEIKKMKATIDPIEK 480
Db 421 VSPDLTREGCKNDVADIVGYKFCDEPGFIPSLIGELITWROEIKKMKATIDPIEK 480
QY 481 MLDYRQARVLT----- 491
Db 481 MLDYRQARVLT----- 491
QY 492 ----- 491
Db 492 ----- 491
QY 539 VKDLKALSFNRETQKSELKVKALIRHRYSGVYSIKLSGRRIKITSGHLSFSYKNGKL 598
Db 539 VKDLKALSFNRETQKSELKVKALIRHRYSGVYSIKLSGRRIKITSGHLSFSYKNGKL 598
QY 599 VKVRGDELPGDLVVPGRKLPESKQVNLVLELLKPREETSNIWMIPYKGRKNFEK 658
Db 599 VKVRGDELPGDLVVPGRKLPESKQVNLVLELLKPREETSNIWMIPYKGRKNFEK 658
QY 659 GMLKTLWIFGEBGERPTAGRYLHLERLGYVKKLRGCEVLDWESLRYRKLLETILN 718
Db 659 GMLKTLWIFGEBGERPTAGRYLHLERLGYVKKLRGCEVLDWESLRYRKLLETILN 718
QY 719 LKYNNSRAVWVEFNSLDVYSIMPIEELKEMIIIGEPGPKIGTFIVDDSPAKLGYI 778
Db 719 LKYNNSRAVWVEFNSLDVYSIMPIEELKEMIIIGEPGPKIGTFIVDDSPAKLGYI 778
QY 779 SSGDEVKDRVFNHSDQVNLLEDIAKLEKFGKVRGGRGYIEVSGKISHAIFRVLADEGR 838
Db 779 SSGDEVKDRVFNHSDQVNLLEDIAKLEKFGKVRGGRGYIEVSGKISHAIFRVLADEGR 838
QY 839 IPELFTSPMDIKVAFLKGLNGMAEELTFSTKSELLVNLILLNSIGVSDIKIEHKGCV 898
Db 839 IPELFTSPMDIKVAFLKGLNGMAEELTFSTKSELLVNLILLNSIGVSDIKIEHKGCV 898
QY 899 YRVYINKKSSNGDVLDSVESIEVEKIEGYVYDLSVDENEMFLVGFGLYAHNSYGY 958
Db 899 YRVYINKKSSNGDVLDSVESIEVEKIEGYVYDLSVDENEMFLVGFGLYAHNSYGY 958
QY 959 GYPAKRWYKSCCAESVTAMGRHYIEMTKEIEKFGFVLYADTDGATATPGEKP-ETI 559
Db 959 GYPAKRWYKSCCAESVTAMGRHYIEMTKEIEKFGFVLYADTDGATATPGEKP-ETI 559
QY 560 KKKAKPEFLKYNSKPLGLELEEGFYLRGFEVAKRYAVIDEBRITTRGLEVVRWMS 619
Db 560 KKKAKPEFLKYNSKPLGLELEEGFYLRGFEVAKRYAVIDEBRITTRGLEVVRWMS 619
QY 1018 KRALEFVDYINSLKPLGLELEEGFYLRGFEVAKRYAVIDEBRITTRGLEVVRWMS 1077
Db 1018 KRALEFVDYINSLKPLGLELEEGFYLRGFEVAKRYAVIDEBRITTRGLEVVRWMS 1077

```

Oy 620 EIAKTEQAKVLEALIKESDSEKAIVEKVDVMEKQVOPLEVIHIOITRNDSEYKAI 679
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1078 EIAKETQARVLEALIKHGNEVAIVVDYETKLNINYEPEKLVITQITRPINEYKAI 1137
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 680 GPHVAIAARLAAKGIKVRGTTISYIVLRGSGKISDRVILLSEDPKKKXDPDYIENQ 739
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1138 GPHVAVAKRLMARGIKRPGWIGYIVLRGDPISRAISIEEPDRKKHXYDAEYIENQ 1197
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 740 VLPAAVLRLLEAGYRKEDLYKSSKQVGLDAMLK 773
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1198 VLPAAERILKAFGRKREDLRMOKTKOVGLGAWIK 1231
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
S71551
DNA-directed DNA polymerase (EC 2.7.7.7) KOD, Intein containing precursor - Pyrococcus s
N:Contains: DNA endonuclease (EC 3.1.-.-) PI-I; DNA endonuclease (EC 3.1.-.-) PI-II; DNA
C:Species: Pyrococcus sp.
A:Variety: Strain KOD1
C:Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000
C:Accession: S71551
      Kakihara, H.; Takagi, M.; Imanaka, T.
      Submitted to the EMBL Data Library, March 1994
      Description: Cloning and analysis of the DNA polymerase gene from a new hyperthermophilic
      A:Reference number: S71551
      A:Accession: S71551
      A:Molecule type: DNA
      A:Residues: 1-1670 <KAK>
      A:Cross-references: EMBL:D29671
      C:Function: <NUCL>
      A:Description: as DNA-directed DNA polymerase, catalyzes the polymerization of DNA at th
      A:Note: DNA-directed DNA polymerase KOD
      C:Function: <EN1>
      A:Description: as DNA endonuclease PI-I, catalyzes the hydrolysis of internal phosphodiester
      A:Note: DNA endonuclease PI-PspI
      C:Function: <EN2>
      A:Description: as DNA endonuclease PI-II, catalyzes the hydrolysis of internal phosphodiester
      A:Note: DNA endonuclease PI-PspII
      C:Superfamily: DNA-directed DNA polymerase KOD
      C:Keywords: endonuclease; hydrolase; nucleotidyltransferase; protein splicing
      F:1-406/Domain: DNA-directed DNA polymerase KOD #status predicted <MA1>
      F:1-406/Domain: DNA-directed DNA polymerase KOD extein 1 #status predicted <XT1>
      F:407-766/Product: DNA endonuclease PI-I (pol KOD Intein 1) #status predicted <MA12>
      F:407-766/Product: DNA endonuclease PI-I (pol KOD Intein 2) #status predicted <XT2>
      F:853-1187/Product: DNA-directed DNA polymerase KOD extein 2 #status predicted <MA13>
      F:853-1187/Product: DNA endonuclease PI-II (pol KOD Intein 2) #status predicted <XT3>
      F:1388-1670/Domain: DNA-directed DNA polymerase KOD extein 3 #status predicted <MA13>
      F:408-767/Cross-Link: peptide (Arg-Ser) #status predicted
      F:851-1388/Cross-Link: peptide (Asn-Ser) #status predicted

Query Match 68.7%; Score 2764.5; DB 2: Length 1670;
Best Local Similarity 35.9%; Pred. No. 2,9e+134;
Matches 599; Conservative 83; Mismatches 88; Indels 899; Gaps 4;

Oy 1 MIFPDYITKDGKPIIRIFKKENGFEKIELDPHFQPIYVALIKDSDAIDEIKALKGERHG 60
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MILPDYITFEDKPIAIRIFKKENGFEKIEYDRTFEPYALLKDDSDAIEVKKITAEKHG 60
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 61 KIVRVADVAKVKKKRLGSDVEYWKLIFFHPOVAPLRLCKIRHPRAVIDIYEDIPFARKY 120
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TVVTVAKRVEKVKKRLGSRVEYWKLIFFTHPOVAPLRLCKIRHGHAVIDIYEDIPFARKY 120
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 121 LIDKGLIPMEGDEELKLAIFDIETFEYHGEDEFGKGEIIMISYADEEARVITWKNIDLPY 180
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LIDKGLIPMEGDEELKLAIFDIQTLIHGSEERABEPIIMISTADEEGARVITWKNVDLPY 180
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 181 VDVVSNREMERIKRVOIVREKDPDLITYNGDNFDLPYLIRAEKLGVTLLGRDKEHE 240
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 VDVVSTEREMIKRFRVAVKEXKDPDLITYNGDNFPAFLAKKRCCKGLNFALGRDGS--E 238
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 241 PIHMGSEFANEIGRIHFDLPVAVRTINLPYTLAEVAVAVGKIKSKLGAETIAI 300
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 PIQRMGDFAEVAVGRHFDLPVIRRTINLPYTLAEVAVAVGQREKVAEETIPA 298
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

QY	301	WEESMUKLAQYSDMOPARATYELGKEFPFMEKLAKLIGOSWVDSRSSTGNLEWYLL	360
Db	299	WETGENLERVARYSMDAKYATYELGKEFPFMEKQSLRLIGOSLMDVSRSSTGNLEWYLL	358
QY	361	RYAYERNELANNKDEDEEYRRRLTTYLGGYVAKESPGRGMENTYLDFFGC	410
Db	359	RKAYERNELANNKDEKELNAR-RQSYEGGYVAKESPGRGMENTYLDFFCHADTKVYVK	417
QY	411	-----	410
Db	418	GKGIINISEVOEGDYVLGIDGMORVKRWEDYKGELVNINGLKCTPNHKLPVYTKNERQ	477
QY	411	-----	410
Db	478	TRIRDSLAKSFLTKKVKGKIITTPLEFELGRATSENIPEEYVLKGELAGILLAECTLLRK	537
QY	411	-----	410
Db	538	DVEYFDSRRKKRRISHQYRVEITIGKDEEFPDRITYIFERLEGTIPSISEKKGNAYTL	597
QY	411	-----	410
Db	598	KYAKKNVYLKKEIMDNIESLHAPSVYLRGFEBDGSYVNRYSIVATOSTKNEMKIKLVS	657
QY	411	-----	410
Db	658	KLLSOLGIPHOTYTYOYOENKDKRSRYLEITGKDGLLFOFLIGFISERNKALLNKAIS	717
QY	411	-----LPSIIIVHN	420
Db	718	ORENNLENNGFYLLSEBPNSTYEYEGKYVDLLECTPYYPFANGILITNHSILPSIIITHN	777
QY	421	VSPDTLEREGCKNVDAPYIVGKFCNDPGFIPISILGELITMROELKKMKATIDPIERK	480
Db	778	VSPDTLEREGCKEYDVAPOVGHRCXDPGFPITSLGLDLEERQIKKKMKATIDPIERK	837
QY	481	MDTYQRVAKLHA- : ::	493
Db	838	LIDYQRAIKILANSILPEWLPVLEBGEVHFVRIGELIDRMWENAGVKREGETEYLE	897
QY	494	-----	493
Db	898	VSGLEVSPFNRTNKAELKRYKALIRHDYSGKYTTIRLKSGRRIKITSGHSLSFVRNGL	957
QY	494	-----	493
Db	958	VEYTGDELPGDLVAVPRLELPERNHVNLVLELLGTPEETLIDIVMTIPVKGRKNFFK	1017
QY	494	-----	493
Db	1018	GMLRTLRKIFQEBEKRPTARRYLRLHLEDLGYVLKIKIGYVLDMDSLKRYRLRYEALYEN	1077
QY	494	-----	493
Db	1078	VRYNGNKRREYLVFENSIRDAVGIMPLKELKEMKIGTLNGFRYRKLIIEVDESLAKLLGYV	1137
QY	494	-----	493
Db	1138	SEGVARKORNPKNQMSYVKVLINEDPEVLDMDERLASRFFGKYVRGRNYVEIPKKIGYLL	1197
QY	494	-----	493
Db	1198	FENMGCVLAENKKRIPEVFTSPKCVRLAFLEGTSSAMASTDEQNALNEKALANQVL	1257
QY	494	-----	493
Db	1258	LNNSYSAVAKIGHDSGYRVYINLELPLVKLDKKKNAYSHVIPREYLVSEYFGKVFQKN	1317
QY	494	-----	493
Db	1318	VSPQTFRMKVEDGLRDPKQORLSWLTIEGDVYLDRAVESVDVEDYDQYVYDLSVEDENENFL	1377


```

      116 FAKRYLLDKGLIPE-----GDEELKLMADI 142
      117 FRRRLDKSLVPEELEEFGCEVDSABSVTTDVRTVEYTRGVOSTGSGANGDLTSPDI 176
      143 ETPYHEG-DEFGKEIIMISYA-DEEARVITMKNIDLPHYDVVSNREMIKRFQVIRE 200
      177 EVRPHGMPDEKDEIIVAGVAGNMVGVSTAGDHLDFEYVEDEHLELFRFELIYD 236
      201 KDPVLTLYNGDNFDPYLIRAEKLVTLIGRDKHEPEKIRHMGDSF--AVEIKRI 258
      237 KKPILVGYNSDNDFPYTRRAAILGALDLGWDGS---KIRMRGRFNATAIKGTV 292
      259 HFDEFPVRRITNPTYLLEAVYEAUIGKTSKIGAEIAIWEESMKLAQVSMDA 318
      293 HVDLPVRRIRMANIDRTILEVYOLGELGEEKIDLPGRMLMEYWDRLDELFRYSLDGV 352
      319 RATYELGKEFPPEAEALAKLIGQSVMDVSRSGTGLVEMYLRAVERNELAPNKDEEE 378
      353 VATHRIEKLILPMLLELTRLVGQPLFDISRMAVGQAEMFLVRAKAYOGELVPNKPSGD 412
      379 YRRRLRTTYLGGYVKEPERGCMENITTYLDFRCLYPSIIVTNVSPDILERECKNYDAP 438
      413 FSSRRGRRAVGGYVKEPEKGHENIVOFDFRSLPSIISKNISPDILTDEESECYAP 472
      439 IVGKFCXDPFGFIPSLIGELITMRKOEIKKKMATIDPIEKKMLDYRORAVKIHAANSYG 498
      473 EYGRFRKSPRGFVPSYIGELISERVRIKEEMKSDDEMEKILWQOEAALKRLANITWG 532
      499 YMGYPKARWYSKECAESVTANGRHVIENTIKEIEKFGKVLVADTDFVATIPG 553
      533 VYGYSRRRWYSMECAEAITANGRDYIKKITAEE-EGFHVYVADTDFVATIPG 586

```

```

RESULT 11
JC7382
DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfurispheera ohwakuenis
N:Alternate names: DNA polymerase alpha, beta, gamma; DNA polymerase I, II, III
C:Species: Sulfurispheera ohwakuenis
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C:Accession: J07382
C:Label: T.; Kurosawa, N.; Itoh, Y. H.; Kimura, N.; Horituchi, T.
DNA Res. 7, 243-251, 2000
A:Title: Sequence analysis of three family B DNA polymerases from the thermoacidophilic
C:Reference number: J07380
A:Accession: J07382
A:Molecule type: DNA
A:Residues: 1-781 <IWA>
A:Cross-references: DDBJ:AB032376
C:Comment: This enzyme has both 3'-5' exonuclease and polymerase activities, and plays a
C:Genetics:
A:Gene: B3
C:Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase

```

```

Query Match          29.9%; Score 1204.5; DB 2; Length 781;
Best Local Similarity 36.6%; Pred. No. 1.7e-54;
Matches 294; Conservative 141; Mismatches 282; Indels 87; Gaps 21;

```

```

      2  IFDPTDITKDGKPIIRIFK-KENGEFKIELDPHFQPIYALLKDSADIEIK-AIKGR 58
      7  ILDESVDVERKPIYIWIVIDKGNRVLLKEKFRPYALVDSYNDIEIRKEIKLISK 66
      59 HGKIVRVADVAVKKKGLGRVYEWKLIFFHPQVPALRGKIREHPAVIDIYEDIPRAK 118
      67 PYSPTSIDV--EKKKFGSVKVLKIEYVPAVVRVYRDEVARIKGVKSVLEADIRPYM 124
      119 RYLLDKGLIPE-----EGDE-----ELKLAFLDIETFYH 147
      125 RYSDIDILKPPYWEALAEVEEIKENNRKVKYELKIKNLKYLEDKIPELKVLAFLIEVY-- 182

```

```

      148 EGDEFG-----KGEIIMISYADEEARVITMKNIDLPHYDVVSNREMIKRFQVIREKD 202
      183 --NKYGSFNRPRDPYIIIGVWTKEGGQELADKYP-----DLRAIREFINVFQITD 231
      203 PDVLTLYNGDNFDPYLIRAEKLVTLIGRDKHEPEKIRHMGDSFAVEIKRIHFDL 262
      232 PDIIIVGINNEDWPLYLERANIRGIRLDVGR-RVNGEPS---QGYYCHYSTIGRLNDL 287
      263 FPPVRRITNPTYLE--AVYEAUIGKTSKIGA-ELIAIWEESMKLAQVSMEDAR 319
      268 YGFAOSIOEVKVTLENADYGLVLPKEKRTIVEMDIPKYWDDEKRDILKYNLDLAK 347
      320 ATYELGKEFPPEAEALAKLIGQSVMDVSRSGTGLVEMYLRAVERNELAPNKDEEEY 379
      348 SAYLLGEVFIPIGIELFTISGLPLDQLASAGHVEMILMEAKIYNLITNK-BERY 406
      360 RRRRLTYLGGYVKEPERGCMENITTYLDFRCLYPSIIVTNVSPDILERECKNYDAP 439
      407 E-----SYEGGLVISPGLGHEEVVLDFFSSMYPSIMIKYNIGPDLVKGECENCWSP- 460
      440 VGYKFCXDPFGFIPSLIGELITMRKOEIKKKMATIDPIEKKMLDYRORAVKIHAANSYG 499
      461 VGHKTRKEPPGLYKNVLEKLIQERKEVKKMKKTIDEDKRVLDKORALKYMANAFYGY 520
      500 MGYPRARWYSKECAESVTANGRHVIENTIKEIEKFGKVLVADTDFVATIPGKPEPI 559
      521 MGLGARWYSKEGAAVYAMGROIISDSAKIAKEK-GFTVIYIGDDISIFVKGGD----- 574
      540 KKKAKEPLKYNISKLPGLLEIEYEGFYLRGPVVA-KKRAYVIDEGCRITTRGLEVVRDM 618
      575 --INSLITEISSKFG--LEIKIDIKYRVFETENKKRAGLIEDGKIDIVGFEAVRGDM 629
      619 SEIAKETAKYLEALIKEDSVKAEIVKDYVEAEIKAOVPEKIVIHQIKDISEKXA 678
      630 CDIAQOVOTWIELILKSGKVEDAIKVKSYFDRIRNFRIEDLITKTNLDEIDV 689
      679 IGPHVAIKRLAAKGIKVRPGTIISYIVLRGSGKISDRY--ILLSEYDPKKHKYDPDY 735
      690 TAPHVAAKKAAGVYLVSKYIGVIVYVKGSGKISDAEPFLVKE---KNKIDVEY 745
      736 IENQVLPAVLRLLEFGYRKEDLK 759
      746 IDKQIIPALRLLEFGVKESSLK 769

```

```

RESULT 12
B56277
DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodicticum occultum
C:Species: Pyrodicticum occultum
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
C:Accession: B56277
C:Label: T.; Ishino, Y.; Doi, H.; Kato, I.
J. Bacteriol. 177, 2164-2177, 1995
A:Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA p
A:Reference number: A56277; MUID:95238290
A:Accession: B56277
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-803 <UEM>
A:Cross-references: GB:D38574; NID:g807829; PIDN:BA07580.1; PID:g807830
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase

```

```

Query Match          29.6%; Score 1190; DB 2; Length 803;
Best Local Similarity 33.7%; Pred. No. 9.8e-54;
Matches 281; Conservative 164; Mismatches 278; Indels 110; Gaps 23;

```

```

      1  MIFPDY-----ITKDGKPIIRIFK-KENGEFKIELDPHFQPIYALLKDSAD--D 45
      8  VLDSSYSLIKKEPVYIIMGTLIDGKRYVL-----LDHRRFPYFALARQYE 55
      46 SAIDEIKAIKGERHGIIVR--VVDAVVKVKKFLGRDVEWKLIFEHPODVPAIRGKIREH 103

```

```

Db 56 DWVEELAA--SIRLSVSKSPITDAKPLDKRFRGPRKAVKITTMIPESVRHVREAVKKI 113
Oy 104 PAVIDIEYDFAKRYLLDKGL-----IPME-----GDEE-----134
Db 114 EGEVDSLEADIFAMRYLIDKRLYPPTYRIFVEDAGRPGRVDRVRYVAGDPEPLADI 173
Oy 135 -----LKLMAFIETFFHEGD-EFGKEIIMISYADEE-BARVITMKNIDLPRVYDVS 185
Db 174 RTIDLPMLVAFIDIEVYSRSGSPNPARPVITVSLRDESGERLIEAGCHD-----225
Oy 166 NREMIKRRVOYVREKDPVLLTYNGDNFDLPYLLKRAKLGVTLLGDKHEPEPKIR 245
Db 226 -DRVLREVEYVRAFDPDIIIVGYSNHPDWPLMERARRLIKIDLVT-RVGAETPTSV 283
Oy 246 MDSFAVEIKGRHFDLPFVVRTIINLPYTLAEVAVLGTKSGAE--IAIWE 302
Db 284 YGH---VSVOGLNVDLDIYAEEMPEIKKTELEVAEYGVCKKSERVITTEMRIPEYD 340
Oy 303 TEESMKKLAQYSMEDARATYELGKEFFPMELAKLIGOSVMDVSRSTGNTLVENYLLRV 362
Db 341 DEKKHQLLERYALDVRATYGLAEKKLPALQISTYGVPLDQVGMVGFRLEMYLMA 400
Oy 363 AYERNELAPNKDEEYRRLRTTYLGIVYKREPRGLMENTIYLDPRCLPSTIYTHNS 422
Db 401 AYDMNELVPRNRE-----RGESEYKGAVALKPLKGVHENVVLDSSMYPSTIMIKYNG 454
Oy 423 PDLT--ERECKKNYD--VAPIGVKFCDFPFISIGELITMROELKKKKA-TIDP 476
Db 455 PRTIYDDPECKYGGCYAPVGVGHFRSPGPFETVLENLKLRLQVKEKKKEPPPS 514
Oy 477 IKKMLDYRQRAVKLHANSYGYMGYPKARWYKSCBASVYAMGRHYIEMTKEIEKFG 536
Db 515 PEYRLTYDEOKALKLANASYGMSHARWYCKRCAEAVTAMGRNLI-LTAIEYARKIG 573
Oy 537 FVLVADNDGFAATIPGEKPEITIKKAKPELYINSKLGLELEGEVLTGGEFV-ARK 595
Db 574 LKVIYIGDSDSLVYVD-----KEKVEKLEFEVEKELD--PEIKIDKTYKKVFEYBAK 624
Oy 596 RYAVIDEGRITTRGLEVVRDMSEIAKETQAKVLEAILKEDSVKEAVEIVDVVEIAK 655
Db 625 RYVGLLEGRIDIVGEAVRGMCELAKEVGEKAEIVLTNGVDAKAIYIREVILQMLE 684
Oy 656 VOYPLEKLYIHQITKDLSEYKAIGPVAIAKRLAKGIVRPGTIIISIVLGRSGKID 715
Db 685 GVPYITKLIMTKLSRIEYEDAPHYMAARMKEAGVSGDVGVIYKSGSVSS 744
Oy 716 RYLLSEYDPKHKYDPOYIENOVLPAYLRILEAGYKEDLKYOSKOVGI 768
Db 745 RAYPYFMVDP--TIDVNYIDHQIVPALRILSYFGTEKOLKAATVQRSL 795

```

RESULT 13

E72515
 Probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: E72515
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum A:Reference number: A72450; MUID:99310339
 A:Accession: E72515
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-784 <KAM>
 A:Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BAAB1109.1; PID:95105797
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2098
 C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 28.3%; Score 1138; DB 2; Length 784;
 Best Local Similarity 34.6%; Pred. No. 4.5e-51;
 Matches 282; Conservative 137; Mismatches 287; Indels 110; Gaps 19;

```

Oy 11 DCKPIRIRKKEGKERKIELDHPQYIYALLKDDSAIDBIKAIKEERKGIYRVYDANK 70
Db 16 DGSRRVVFY---GEPR-----PYFVLPDGSVGLQDLAMIRRLSPSSPILSV 63
Oy 71 VKKFLGRVVEWKLFEPHODVPALRGKIREHPAVIDIYEDIPAKRYLIDKGLPM- 129
Db 64 VRRRFIGREVEALKVTLVPASVREYREAVRRLGVRDVLADIPALRIFIIDFNLYPM 123
Oy 130 -----EGD-----EELKMAFIETFFYH-EGDE 152
Db 124 WYVAEVEVAVPHGVSVDRAYTLSDIREDETRIGEDPLKGLVMAFDIEVYSKMTPP 183
Oy 153 GKGEIIMISYD-----EEBARIYTKNIDLPRVYDVSNREMIKRRVOYVREKDP 204
Db 184 KDDPVIMIGLQAGSEIILEADR-----SDKKYIAGFEVVRKSIDP 227
Oy 205 VLTYYNGDNFDLPYLLKRAKLGVTLLGDKHEPEPKIRHMGDSPAVEIKGRHFDLP 264
Db 228 VIVGYQNPFDMYLYERARVGVKLAAGRVSVEPQGLYG-----HYSYGLNVDLD 282
Oy 265 VVRTIINLPYTLAEVAVLGTK--SKIGAE--EIAIWEESMKKLAQYSMEDARA 320
Db 283 FAEELHEVAVKLTLEAVD-YLGVVKIGERYTLEMMQIGEWMDPSKREILRKLRDVRS 341
Oy 321 TYELGKEFFPMELAKLIGOSVMDVSRSTGVLVEMYLILRVAYENELAPNKPDDEEVR 380
Db 342 TWGLAKFLPFGAELSOVGLPLDQVMAASVGRMLRLREAKIGELVAPNVERSEGR 401
Oy 381 RRLTYTLGGVYKPEPRGLMENTIYLDPRCLYPSIIVTHNSDPTLEREGCKNYD-----435
Db 402 -----YACAIYLRPKRGVHEDIAVLDFAISMVNIWAKNVGGDTLVRG--EEYSGEEVY 454
Oy 436 VAPIGVKFCDFPFISIGELITMROELKKKKA-TIDPIEKKMLDYRQRAVKLAN 494
Db 455 TAPEVGHKFRKSPGPFKKILERFLSMRQIRSEMKHPPDSPEYKILDEROKAAILLAN 514
Oy 495 SYVGYGPKARWYKSCBASVYAMGRHYIEMTKEIEKFGKGVYUADPGFYATIPGE 554
Db 515 ASYGYGMFHARWYKSCBASVYAMGRSIIITRAIRAGF-LGLEVYIGDTSLFVKNDE 573
Oy 555 KEPTIKKAKPELYINSKLGLELEGEVLTGGEFV-AKKRYAVIDEGRITTRGLEV 613
Db 574 KYERLLRFVEBELG-----DIKVDKYVRVFEFFTEAKKRYVGLYDGKIDVVGFE 624
Oy 614 VRDMSEIAKETQAKVLEAILKEDSVKEAVEIKVDVVEIAKQVPLEKLYIHQITKDL 673
Db 625 VRGDMSELAKETQFKVAEIVLKTGSVDEADVVRNITELKIRGOVDMRKLVIYKLTTRPP 684
Oy 674 SEYKAIGPVAIAKRLAAGIKVRPGTIIISIVLGRSGKISDQVILLSEYDPKHKHYDDP 733
Db 685 SMIEARQPHVTALLMERAGIKVEPAKIGIVYTKSGSPLYTRA--KPYFMASKEEYDVE 742
Oy 734 YYIENQVLPAYLRILEAGYKEDLKYOSKOVGLD 769
Db 743 YYVDKQVYVPAALRIIDYFGYTEKRLKGGGQSYLLD 778

```

RESULT 14

T05731
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - soybean
 C:Species: Glycine max (soybean)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 18-Jun-1999
 C:Accession: T05731
 R:Collins, J.T.B.; Cannon, G.C.; Heinhorst, S.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: 215439
 A:Accession: T05731
 A:Status: preliminary; translated from GB/EMBL/DDBJ

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:17:02 ; Search time 14.68 Seconds
(Without alignments)
1287.834 Million cell updates/sec

Title: US-09-803-165-34
Perfect score: 4026
Sequence: 1 MIFPDYIKDKGPIIRIFK.....KEDLYQSSKQVGLDMLKK 774

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3225.5	80.1	774	3	US-08-902-632-2
2	3222.5	80.0	774	3	US-08-073-354-1
3	3222.5	80.0	774	3	US-08-656-005A-1
4	3222.5	80.0	774	4	US-09-073-259-1
5	3222.5	80.0	774	4	US-09-363-095-1
6	3222.5	80.0	774	4	US-09-418-027-1
7	3209.5	79.7	778	2	US-08-906-925-4
8	3151	78.3	775	1	US-07-966-278-1
9	3151	78.3	775	1	US-08-424-921-1
10	3151	78.3	775	2	US-08-556-355A-1
11	3151	78.3	775	2	US-07-803-627A-1
12	3151	78.3	776	2	US-08-688-649-37
13	3139	78.0	779	1	US-08-375-134-12
14	3139	78.0	779	5	PCT-US95-15263-12
15	2502.5	62.2	1022	1	US-08-271-364A-8
16	2502.5	62.2	1022	2	US-08-222-715B-27
17	2069	51.4	1019	1	US-08-271-364A-7
18	2069	51.4	1019	2	US-08-232-715B-26
19	1595	39.6	366	1	US-08-239-284A-2
20	1403	34.8	788	2	US-08-907-166-6
21	1194	29.7	803	2	US-08-907-166-4
22	1191	29.6	803	1	US-08-062-368-4
23	1180	29.3	803	1	US-08-062-368-2
24	611	15.2	1107	1	US-08-366-577-2
25	591	14.7	877	5	PCT-US96-00005-2
26	591	14.7	877	5	PCT-US96-00005-2
27	588.5	14.6	762	2	US-08-907-166-8
					Sequence 10, Appl

28	585.5	14.5	1462	3	US-07-792-600-31	Sequence 31, Appl
29	585.5	14.5	1462	3	US-09-157-021-31	Sequence 31, Appl
30	585.5	14.5	1462	3	US-09-156-842-31	Sequence 31, Appl
31	479.5	11.9	1009	2	US-08-680-326-31	Sequence 6, Appl
32	467	11.6	783	1	US-08-101-593-6	Sequence 6, Appl
33	467	11.6	783	1	US-08-465-995A-6	Sequence 6, Appl
34	467	11.6	783	2	US-08-465-994C-6	Sequence 32, Appl
35	463.5	11.5	1015	2	US-08-680-326-32	Sequence 117, App
36	463	11.5	837	2	US-08-680-326-117	Sequence 34, Appl
37	461	11.5	1012	2	US-08-680-326-34	Sequence 40, Appl
38	458	11.4	1094	2	US-08-680-326-40	Sequence 10, Appl
39	457	11.4	108	3	US-09-073-354-10	Sequence 10, Appl
40	457	11.4	108	3	US-08-656-005A-10	Sequence 10, Appl
41	457	11.4	108	4	US-09-073-259-10	Sequence 10, Appl
42	457	11.4	108	4	US-09-363-095-10	Sequence 10, Appl
43	457	11.4	108	4	US-09-418-027-10	Sequence 10, Appl
44	433.5	10.8	1008	2	US-08-680-326-30	Sequence 30, Appl
45	426.5	10.6	1097	2	US-08-680-326-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1
US-08-902-632-2
Sequence 2, Application
Patent No. 6008025
GENERAL INFORMATION
APPLICANT: KOMATSUBARA, Hideyuki
APPLICANT: KITABAYASHI, Masao
APPLICANT: KAMIMURA, Hideki
APPLICANT: KAWAKAMI, Bunsei
APPLICANT: KAWAMURA, Yoshihisa
APPLICANT: TAKAGI, Masahiro
APPLICANT: IMANAKA, Tadayuki
TITLE OF INVENTION: Modified Thermostable DNA Polymerase,
TITLE OF INVENTION: and DNA Polymerase Composition for Nucleic Acid
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1 Broadway
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10004
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 Inch floppy disk, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/902,632
FILING DATE: Concurrent Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 198911/96
FILING DATE: 29-JUL-1996
APPLICATION NUMBER: JP 200446/96
APPLICATION NUMBER: 30-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: GREASON, Edward W.
REGISTRATION NUMBER: 18,918
REFERENCE/DOCKET NUMBER: 2418/7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-425-7200
TELEFAX: 212-425-5288
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein


```

QY 421 VSPDTEREGCKNDVAVIVGKCKDPGFIPIISLIGLITMRQEIKKMKMATIDPIEKK 480
Db 418 VSPDTEREGCKNDVAVIVGKCKDPGFIPIISLIGLITMRQEIKKMKMATIDPIEKK 477
QY 481 MLDYRQRAVKRKAHSYGYMGYPRKAWYKSCCAESVTAMGNHYEMTKEIEEKGFEVYL 540
Db 478 LLDYRQRAIKILANSYGYGYARAWYCKCAESVTAMGNHYEMTKEIEEKGFEVYL 537
QY 541 YADIDGFATIPGADAEIVYKKAAMEFLNYINAKLPALAELEYEGFYKKGFEVYTKKRAYI 600
Db 538 KSDIDGFATIPGADAEIVYKKAAMEFLNYINAKLPALAELEYEGFYKKGFEVYTKKRAYI 597
QY 601 DEGRITTRGLEVVRDMSEIAKETQAKVLEAILKEDSEKAVELVKDVEIAKAYOVL 660
Db 598 DEGRITTRGLEVVRDMSEIAKETQAKVLEAILKEDSEKAVELVKDVEIAKAYOVL 657
QY 661 EKVYIHEDITDLSEYKAIQHVVAIAKRLAKGKIKVRGTTIISYIVLNGSKISDRVLL 720
Db 658 EKVYIHEDITDLSEYKAIQHVVAIAKRLAKGKIKVRGTTIISYIVLNGSKISDRVLL 717
QY 721 SEYDPKHKHYDPDYIENOVPAVLRIIEAFGYRKEDLKYOSKQVGLDAMLK 773
Db 718 DEFDPTRKHKYAEYIENOVPAVLRIIEAFGYRKEDLKYOSKQVGLDAMLK 770

RESULT 3
US-08-656-005A-1
: Sequence 1, Application US/08656005A
: Patent No. 6054301
: GENERAL INFORMATION:
: APPLICANT: KITABAYASHI, Masao
: APPLICANT: ARAKAWA, Taku
: APPLICANT: INOUE, Hiroaki
: APPLICANT: KAWAKAMI, Bunsei
: APPLICANT: KAWAMURA, Yoshihisa
: APPLICANT: IMANAKA, Tadayuki
: APPLICANT: TAKAGI, Masahiro
: APPLICANT: MORIKAWA, Masaaki
: TITLE OF INVENTION: A Method of Amplifying Nucleic
: TITLE OF INVENTION: Acid and A Reagent Therefor
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenyon & Kenyon
: STREET: 1025 Connecticut Avenue, N.W., Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WordPerfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/656, 005A
: FILING DATE: 24 MAY 1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 134096/95
: FILING DATE: 31 MAY 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Toffenetti, Judith L.
: REGISTRATION NUMBER: 39,048
: REFERENCE/DOCKET NUMBER: 2418/3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-429-1776
: TELEFAX: 202-429-0796
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 774 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: linear

```

```

: MOLECULE TYPE: protein
: US-08-656-005A-1
Query Match      80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred. No. 6,56-243;
Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

QY 1 MIFPDYITKCKPIRIIRKENGKFEKTELDPHOPYIYALLKDSADLKAIGKEHG 60
Db 1 MILDYITTEGCKPIYIRLFKENGKFEKTELDPEYFVYALLKDSADLKAIGKEHG 60
QY 61 KIVRVDAVAKKKKFLGSDVEMWKLIFEPHODVPALRCKIENHPAVIDIYEYDIPFARY 120
Db 61 TVVIVYKREKQKKNFLGSPVEWKLIFEPHODVPALRCKIENHPAVIDIYEYDIPFARY 120
QY 121 LIDKGLIPMEGDEBELKMAFDIETFPYHEGDEFGKEIIMISYADEEARVITWKNDLPY 180
Db 121 LIDKGLVPMEGDEBELKMAFDIOTLYHEGSEFAEGPIIMISYADEEGARVITWKNDLPY 180
QY 181 VDVSNEREMTKRFQIYREKDPDVLITVNGDNFDPYLIRAEKLGVTLLGRDKEHP 240
Db 181 VDVSSTEREMTKRDLRVYKEDPDVLITYNGDNFDPYLRKCEKLGINFALGRDGS--E 238
QY 241 PKIRMGDSFAVEIKGRIFHDFPVVRRITNLPTTYLAEVAVLGGTKSKLGAEEIAI 300
Db 239 PKIRMGDSFAVEYKGRIFHDFPVVRRITNLPTTYLAEVAVLGGTKSKLGAEEIIPA 298
QY 301 WETESMKKLAOYSMEDARATYELGKEFPMEAEIAKLGOSVWDVSRSSGNTLYEWYLL 360
Db 299 WETENLENLVARYSMEDAKVYELGKEFLPMEAOQLSRIGOSLMDVSRSSGNTLYEWYLL 358
QY 361 RVAYERNLAPNKPDEEYRRRLRTTYLGGYVYKPEBERGLMNITVYLDRCYPSIIVHN 420
Db 359 RKATERNELAPNKPDEEYLRAR-RQSYEGYVYKPEBERGLMNITVYLDRCYPSIIVHN 417
QY 421 VSPDTEREGCKNDVAVIVGKCKDPGFIPIISLIGLITMRQEIKKMKMATIDPIEKK 480
Db 418 VSPDTEREGCKNDVAVIVGKCKDPGFIPIISLIGLITMRQEIKKMKMATIDPIEKK 477
QY 481 MLDYRQRAVKRKAHSYGYMGYPRKAWYKSCCAESVTAMGNHYEMTKEIEEKGFEVYL 540
Db 478 LLDYRQRAIKILANSYGYGYARAWYCKCAESVTAMGNHYEMTKEIEEKGFEVYL 537
QY 541 YADIDGFATIPGADAEIVYKKAAMEFLNYINAKLPALAELEYEGFYKKGFEVYTKKRAYI 600
Db 538 KSDIDGFATIPGADAEIVYKKAAMEFLNYINAKLPALAELEYEGFYKKGFEVYTKKRAYI 597
QY 601 DEGRITTRGLEVVRDMSEIAKETQAKVLEAILKEDSEKAVELVKDVEIAKAYOVL 660
Db 598 DEGRITTRGLEVVRDMSEIAKETQAKVLEAILKEDSEKAVELVKDVEIAKAYOVL 657
QY 661 EKVYIHEDITDLSEYKAIQHVVAIAKRLAKGKIKVRGTTIISYIVLNGSKISDRVLL 720
Db 658 EKVYIHEDITDLSEYKAIQHVVAIAKRLAKGKIKVRGTTIISYIVLNGSKISDRVLL 717
QY 721 SEYDPKHKHYDPDYIENOVPAVLRIIEAFGYRKEDLKYOSKQVGLDAMLK 773
Db 718 DEFDPTRKHKYAEYIENOVPAVLRIIEAFGYRKEDLKYOSKQVGLDAMLK 770

RESULT 4
US-09-073-259-1
: Sequence 1, Application US/09073259
: Patent No. 6143536
: GENERAL INFORMATION:
: APPLICANT: IMANAKA, Tadayuki
: APPLICANT: TAKAGI, Masahiro
: APPLICANT: MORIKAWA, Masaaki
: TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenyon & Kenyon

```



```

Db 299 WETGENLERVARYSMEDAKVTEELGKEFLPMEQLSRLIGSLMDVSRSSSTGNLWEMFL 358
OY 361 RFAVERNELANRKPDEEYRRRLRTTYLGGVYKPERGEMNTINYLDPRCLYPSIIYTHN 420
Db 359 RKAVERNELANRKPDEKELARR-RQSYEGGYVKEPERGLMNTINYLDPRSLYPSIIYTHN 417
OY 421 VSPDTLREGCKNDVAPIVYGKFCDFPGFIPSLIGELITMROEIKKKMATIDPLEKK 480
Db 418 VSPDTLNRGCKEYDVAPOVGHRCCKDPGFPISLIGDLLEERQIKKKMATIDPLEKK 477
OY 481 MLDYRORAVKLHANSYGYMGYPKARWYKSCAASVTAMGRHYIEMTIKEIEKFGFVYL 540
Db 478 LLDYRORAIKILANSYGYGYARARWYKSCAASVTAMGREYITMTIKEIEKFGFVYI 537
OY 541 YADTDGFATIPGKPEPIKKAKEFLKYINSKLPGLLEIEYBEGFYLRGFFVAKKRAVI 600
Db 538 YSDTDGFATIPGADAEVTKKAMEFLNYINAKLPGLALEIEYBEGFYLRGFFVTKKRAVI 597
OY 601 DEGRITTRGLEVVRDMSSEIAKETQAVLEAILKEDSVEKAVEIVKDVEIARAYVPL 660
Db 598 DEGRITTRGLEIYRRDMSSEIAKETQAVLEAILKDGVEKAVRIVKEVTEKLSKEYVRP 657
OY 661 EKLVIHQITDLSYKAIQPHVAIAKRLAKGIVRPGTIISYIVLRGSKISDRVILL 720
Db 658 EKLVIHQITDLDKYATGPHVAIVAKRLARGVKIRPGTIISYIVLRGSGRIDRAIPF 717
OY 721 SEYPPKHKYDPDYIENQVLPVLRILAEAGYRKREDKYSQVGLDAMLK 773
Db 718 DEFDPHKHYDAEYIENQVLPVLRILAEAGYRKREDLRYOKTROVGLSAMLK 770

```

RESULT 6

US-09-418-027-1

Sequence 1, Application US/09418027

Patent No. 6225065

GENERAL INFORMATION:

APPLICANT: KITABAYASHI, Masao

APPLICANT: ARAKAWA, Taku

APPLICANT: INOUE, Hiroaki

APPLICANT: KAWAKAMI, Bunsei

APPLICANT: KAWAMURA, Yoshihisa

APPLICANT: IMANAKA, Tadayuki

APPLICANT: TAKAGI, Masahiro

APPLICANT: MORIKAWA, Masaki

TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 6.1 Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/418, 027

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/073, 354

FILING DATE:

APPLICATION NUMBER: JP 134096/95

FILING DATE: 31 MAY 1995

ATTORNEY/AGENT INFORMATION:

NAME: Toffenetti, Judith L.

REGISTRATION NUMBER: 39, 048

REFERENCE/DOCKET NUMBER: 2418/9

TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: 202-429-1776
: TELEFAX: 202-429-0796
: INFORMATION FOR SRO ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 774 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-418-027-1

```

Query Match

80.0%; Score 3222.5; DB 4; Length 774;

Best Local Similarity 77.4%; Pred. No. 6,5e-243;

Matches 598; Conservativity 83; Mismatches 89; Indels 3; Gaps 2;

```

OY 1 MIFDTYITKDKPRIIRIFKKNCEFEIELDPHFOPYIYALLKDDSAIDEIKAIKGRHG 60
Db 1 MIFDTYITKDKPRIIRIFKKNCEFEIELDPHFOPYIYALLKDDSAIDEIKAIKGRHG 60
OY 61 KIYVNVDAVKKKKKFLGRDVEVMKLIFFENPODVALNGKITEHRAVIDIYEDIPFAKY 120
Db 61 TVVTIVKRVKQKKFLGRDVEVMKLIFFENPODVALNGKITEHRAVIDIYEDIPFAKY 120
OY 121 LIDGLIPMEGDEBLKIMAFDIETFFYHGDGDFGKEIIMISYADEBARVITMKNIDLPY 180
Db 121 LIDGLIPMEGDEBLKIMAFDIETFFYHGDGDFGKEIIMISYADEBARVITMKNIDLPY 180
OY 181 VDVSNSNERMKRRVOIVREKDPVLTLYNGDNFDLPYLIRAKLGVTLLGDKHEPE 240
Db 181 VDVSNSNERMKRRVOIVREKDPVLTLYNGDNFDLPYLIRAKLGVTLLGDKHEPE 240
OY 241 PIHRMGDSFAVEIKGRHFDLPFVVRRTINLPYTLAEVYEAVALGKTSKLGAEITAAI 300
Db 241 PIHRMGDSFAVEIKGRHFDLPFVVRRTINLPYTLAEVYEAVALGKTSKLGAEITAAI 300
OY 299 PKIQMGDRFAVEYKGRHFDLPYIRRTINLPYTLAEVYEAVALGKTSKLGAEITAAI 360
Db 299 PKIQMGDRFAVEYKGRHFDLPYIRRTINLPYTLAEVYEAVALGKTSKLGAEITAAI 360
OY 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEIAKLIGSVMDVSRSSTGNLWEMYL 360
Db 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEIAKLIGSVMDVSRSSTGNLWEMYL 360
OY 359 RKAVERNELANRKPDEKELARR-RQSYEGGYVKEPERGLMNTINYLDPRSLYPSIIYTHN 417
Db 359 RKAVERNELANRKPDEKELARR-RQSYEGGYVKEPERGLMNTINYLDPRSLYPSIIYTHN 417
OY 421 VSPDTLREGCKNDVAPIVYGKFCDFPGFIPSLIGELITMROEIKKKMATIDPLEKK 480
Db 421 VSPDTLREGCKNDVAPIVYGKFCDFPGFIPSLIGELITMROEIKKKMATIDPLEKK 480
OY 418 VSPDTLNRGCKEYDVAPOVGHRCCKDPGFPISLIGDLLEERQIKKKMATIDPLEKK 477
Db 418 VSPDTLNRGCKEYDVAPOVGHRCCKDPGFPISLIGDLLEERQIKKKMATIDPLEKK 477
OY 481 MLDYRORAVKLHANSYGYMGYPKARWYKSCAASVTAMGRHYIEMTIKEIEKFGFVYL 540
Db 481 MLDYRORAVKLHANSYGYMGYPKARWYKSCAASVTAMGRHYIEMTIKEIEKFGFVYL 540
OY 478 LLDYRORAIKILANSYGYGYARARWYKSCAASVTAMGREYITMTIKEIEKFGFVYI 537
Db 478 LLDYRORAIKILANSYGYGYARARWYKSCAASVTAMGREYITMTIKEIEKFGFVYI 537
OY 541 YADTDGFATIPGKPEPIKKAKEFLKYINSKLPGLLEIEYBEGFYLRGFFVAKKRAVI 600
Db 541 YADTDGFATIPGKPEPIKKAKEFLKYINSKLPGLLEIEYBEGFYLRGFFVAKKRAVI 600
OY 538 YSDTDGFATIPGADAEVTKKAMEFLNYINAKLPGLALEIEYBEGFYLRGFFVTKKRAVI 597
Db 538 YSDTDGFATIPGADAEVTKKAMEFLNYINAKLPGLALEIEYBEGFYLRGFFVTKKRAVI 597
OY 601 DEGRITTRGLEVVRDMSSEIAKETQAVLEAILKEDSVEKAVEIVKDVEIARAYVPL 660
Db 601 DEGRITTRGLEVVRDMSSEIAKETQAVLEAILKEDSVEKAVEIVKDVEIARAYVPL 660
OY 598 DEGRITTRGLEIYRRDMSSEIAKETQAVLEAILKDGVEKAVRIVKEVTEKLSKEYVRP 657
Db 598 DEGRITTRGLEIYRRDMSSEIAKETQAVLEAILKDGVEKAVRIVKEVTEKLSKEYVRP 657
OY 661 EKLVIHQITDLSYKAIQPHVAIAKRLAKGIVRPGTIISYIVLRGSKISDRVILL 720
Db 661 EKLVIHQITDLSYKAIQPHVAIAKRLAKGIVRPGTIISYIVLRGSKISDRVILL 720
OY 658 EKLVIHQITDLDKYATGPHVAIVAKRLARGVKIRPGTIISYIVLRGSGRIDRAIPF 717
Db 658 EKLVIHQITDLDKYATGPHVAIVAKRLARGVKIRPGTIISYIVLRGSGRIDRAIPF 717
OY 721 SEYPPKHKYDPDYIENQVLPVLRILAEAGYRKREDKYSQVGLDAMLK 773
Db 721 SEYPPKHKYDPDYIENQVLPVLRILAEAGYRKREDKYSQVGLDAMLK 773
OY 718 DEFDPHKHYDAEYIENQVLPVLRILAEAGYRKREDLRYOKTROVGLSAMLK 770
Db 718 DEFDPHKHYDAEYIENQVLPVLRILAEAGYRKREDLRYOKTROVGLSAMLK 770

```

RESULT 7

US-08-906-925-4

Sequence 4, Application US/08906925

Patent No. 5882904

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

```

Oy 1 MIFPDYITTKGKPIIRFKKENGKFKLEDPHPQPIYALLKQDSADLKAIGKGEHG 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MILDVDYITEGKRVIRLFKKENKFKIEHDRTPRPIYALLRQDSKIEEKKITGERHG 60
Oy 61 KIVRVDAVKKKKFLGNDVEMKLIIEHPDVPALRCKIREHNAVIDIYEDYIPAKRY 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 KIVRVDAVKKKKFLGNDVEMKLIIEHPDVPALRCKIREHNAVIDIYEDYIPAKRY 120
Oy 121 LIDGLIPMEGDEBLKMAFDIETFFYHGEDEFGKEIIMISYADEBEARVITMKNIDLPY 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 LIDGLIPMEGDEBLKMAFDIETFFYHGEDEFGKEIIMISYADEBEARVITMKNIDLPY 180
Oy 181 VDVVSNREMIKKRFVOIYREKDPVLITNGDNFDLPYLIRAKKLGITLLGDKHEPE 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 VDVVSNREMIKKRFVOIYREKDPVLITNGDNFDLPYLIRAKKLGITLLGDKHEPE 240
Oy 241 PKIRMGDSFAVEIKRIFHFDLPVVRRTINLPYTLAEVAVLIGTKSKLGAETIAAI 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 239 PKMORIGDMTAVEVKGRIHFDLVHIRTINLPYTLAEVAVLIGTKSKLGAETIAAI 300
Oy 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEALAKIGOSVWDVSRSSSTGNLVEMYL 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 299 WESGENSELRAVAKYSMEDAKATYELGKEFLPMEIOLSRVGPMDVSRSSSTGNLVEMYL 360
Oy 361 RVAERNEMLAPNKPDEEYRRLRTTYLIGVYKPEPENGIMNTIYLDPRCLYPSIIYTHN 420
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 359 RKAVERNEVAPNKPSEETQORLRESYTGFGVKEPEKGLMNIYIYLDPRALYPSIIITHN 418
Oy 421 VSPDTLEREGCKNDVAPIVYKFCDFPGFIPSIILGELITMROEIKKMKATIDPLEKK 480
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 419 VSPDTLMEGCKNDIAPOVGHKFCDFPGFIPSIILGELITMROEIKKMKATIDPLEKI 478
Oy 481 MLDYRORAVKLIANSYGYMGYPKARWYSKECAESVTAMGHRHYTEMTIKETEEKFGFKVL 540
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 479 LLDYRORAVKLIANSYGYMGYPKARWYSKECAESVTAMGHRHYTEMTIKETEEKFGFKVL 538
Oy 541 YADYDGFATIPGEPETIKKAKKEFLYKINSKLPGLLEEGFYLGPFVAKKRYAVI 600
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 539 YIDYDGLATIPGSESEIKKALEFVYKINSKLPGLLEEGFYLGPFVAKKRYAVI 598
Oy 601 DEGRITTRGLEVVRWOMSEIAKETQAKVLEAILKEDSVKAEVAVIYKDVVEIAKYQVPL 660
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 599 DEEKVITRGLLEIYRDMSEIAKETQAVLEETILKHGVEEAVRIYEVIOKLANYEIRP 658
Oy 661 EKLVIHEITDLSERYKAGHPVAVAKRLAKGIKVRPGTIIISYIVLRGSKISIDRVILL 720
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 659 EKLVIHEITRPLHEYKAGHPVAVAKRLAKGIVKGVAVIYIVLRGSDGPISNRALIA 718
Oy 721 SEYDPKKHKYDPDYIENQVPAVLRIEAGYRKREDKLYOSSKOVGLDAML 772
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 719 SEYDPKKHKYDPDYIENQVPAVLRIEAGYRKREDKLYOSSKOVGLDAML 770

```

Query Match 78.3%; Score 3151; DB 1; Length 775;
 Best Local Similarity 74.6%; Pred. No. 2.4e-237;
 Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

```

Oy 1 MIFPDYITTKGKPIIRFKKENGKFKLEDPHPQPIYALLKQDSADLKAIGKGEHG 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MILDVDYITEGKRVIRLFKKENKFKIEHDRTPRPIYALLRQDSKIEEKKITGERHG 60
Oy 61 KIVRVDAVKKKKFLGNDVEMKLIIEHPDVPALRCKIREHNAVIDIYEDYIPAKRY 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 KIVRVDAVKKKKFLGNDVEMKLIIEHPDVPALRCKIREHNAVIDIYEDYIPAKRY 120
Oy 121 LIDGLIPMEGDEBLKMAFDIETFFYHGEDEFGKEIIMISYADEBEARVITMKNIDLPY 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 LIDGLIPMEGDEBLKMAFDIETFFYHGEDEFGKEIIMISYADEBEARVITMKNIDLPY 180
Oy 181 VDVVSNREMIKKRFVOIYREKDPVLITNGDNFDLPYLIRAKKLGITLLGDKHEPE 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 VDVVSNREMIKKRFVOIYREKDPVLITNGDNFDLPYLIRAKKLGITLLGDKHEPE 240
Oy 241 PKIRMGDSFAVEIKRIFHFDLPVVRRTINLPYTLAEVAVLIGTKSKLGAETIAAI 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 239 PKMORIGDMTAVEVKGRIHFDLVHIRTINLPYTLAEVAVLIGTKSKLGAETIAAI 300
Oy 301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEALAKIGOSVWDVSRSSSTGNLVEMYL 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 299 WESGENSELRAVAKYSMEDAKATYELGKEFLPMEIOLSRVGPMDVSRSSSTGNLVEMYL 360
Oy 361 RVAERNEMLAPNKPDEEYRRLRTTYLIGVYKPEPENGIMNTIYLDPRCLYPSIIYTHN 420
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 359 RKAVERNEVAPNKPSEETQORLRESYTGFGVKEPEKGLMNIYIYLDPRALYPSIIITHN 418
Oy 421 VSPDTLEREGCKNDVAPIVYKFCDFPGFIPSIILGELITMROEIKKMKATIDPLEKK 480
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 419 VSPDTLMEGCKNDIAPOVGHKFCDFPGFIPSIILGELITMROEIKKMKATIDPLEKI 478
Oy 481 MLDYRORAVKLIANSYGYMGYPKARWYSKECAESVTAMGHRHYTEMTIKETEEKFGFKVL 540

```

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/424,921
 FILING DATE: 19-Apr-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/803,627
 FILING DATE: 02-DEC-1991
 APPLICATION NUMBER: US 07/620,568
 FILING DATE: 03-DEC-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,073
 FILING DATE: 19-FEB-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/776,552
 FILING DATE: 15-OCT-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Fitting, Thomas
 REGISTRATION NUMBER: 34,163
 REFERENCE/DOCKET NUMBER: STG0100P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-792-3680
 TELEFAX: 619-792-8477
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 775 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-424-921-1

```

DB 479 LLDYRQKAKRLANSFYGYGAKAMWCKECAESTYANGRKYIELVWKELEBEKGFVYL 538
OY 541 YADTDGFYATIPGEEKPETIKKKAKKEFLKYINSKLPGLLEVEGEFVLRGFFPAKKRYAVI 600
DB 539 YIDTDGLVATITPGEGSEELKKALEFVKYINSKLPGLLEVEGEFVLRGFFPAKKRYAVI 598
OY 601 DEGRITTRGLEVVRDMSSEIAKETQAVLEAILKEDSVKAVEIVKQVVEIAKYQVPL 660
DB 599 DEEGKVTTRGLEIVRDMSEIAKETQAVLEITLKHGVEEAVRIYKIOKLANYEIRP 658
OY 661 EKLVIHEOITKDSYKAGPHVAIAKRLAAGKIRPGTIISYILRSGKISDPAVL 720
DB 659 EKLAIVEOTTRPLHEKKAIPHAAKLAAGVKIKPGMIGVILYLRGDPISRAILTA 718
OY 721 SEYDPKHKHYDDPYIENOVLPVLRILEAFGRKEDLYQSSQOYGLDAML 772
DB 719 EEYDPKHKHYDAEYIENOVLPVLRILEGFGRKEDLYQKTRQVGLTSM 770

```

RESULT 10

US-08-556-355A-1

Sequence 1, Application US/08556355A

Patent No. 5866395

GENERAL INFORMATION:

APPLICANT: MATHUR, Eric J.

TITLE OF INVENTION: Purified Thermostable Pyrococcus

TITLE OF INVENTION: furiosus DNA Polymerase I

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESS: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/556,355A

FILING DATE: 13-NOV-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/424,921

FILING DATE: 19-APR-1995

- PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/803,627

FILING DATE: 02-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/779,846

FILING DATE: 21-OCT-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/657,073

FILING DATE: 19-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/620,568

FILING DATE: 03-DEC-1990

ATTORNEY/AGENT INFORMATION:

NAME: Barker, M. Paul

REGISTRATION NUMBER: 32,013

REFERENCE/DOCKET NUMBER: 04121, 0004-02

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)408-4000

TELEFAX: (202)408-4400

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 775 amino acids

TYPE: amino acid

STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-556-355A-1

```

Query Match 78.3%; Score 3151; DB 2; Length 775;

Best Local Similarity 74.6%; Pred. No. 2,4e-237;

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

```

OY 1 MFEDDYTKCKGPIIRIFKKENGFEKTELDHPQPIYVALLKDDSAIDEIKAGERHG 60
DB 1 MLDVDYITEEKKPIYRIFKKENGFKIEHDTFFPIYALLRDSKIEYVKITGERHG 60
OY 61 KIVRVADVAKKKKFLGRDVEYWKILFEHPDVPALRGKIREPAVIDIYEYDIPARKY 120
DB 61 KIVRVADVAKKKKFLGRDVEYWKILFEHPDVPALRGKIREPAVIDIYEYDIPARKY 120
OY 121 LIDKGLIPMEGDEELKLAFLDLETYHGGDEGCKEITIMISYADEEAKVITWKIDLPY 180
DB 121 LIDKGLIPMEGDEELKLAFLDLETYHGGDEGCKEITIMISYADEEAKVITWKIDLPY 180
OY 181 VDVVNEREMIKRFYOIVREKDPDVLITYGNDPFLPYLIRKAEKLGVTLLGRDKEHE 240
DB 181 VDVVNEREMIKRFYOIVREKDPDVLITYGNDPFLPYLIRKAEKLGVTLLGRDKEHE 240
OY 239 PKMQIGDMTAVEVGRHIFDLHYVITRTINLPYTLVAVEAIGKPREKRYADEIAKA 298
DB 239 PKMQIGDMTAVEVGRHIFDLHYVITRTINLPYTLVAVEAIGKPREKRYADEIAKA 298
OY 301 WTEESMKKLAQYSMEDARATYELKEFFPMAPAEALAKLIGQSVMPVSSSGNLYEWYLL 360
DB 301 WTEESMKKLAQYSMEDARATYELKEFFPMAPAEALAKLIGQSVMPVSSSGNLYEWYLL 360
OY 359 WESGENSELKVAYVSMEDAKATYELKEFLPMETQLSRVQGLPMDVSSSGNLYEWYLL 358
DB 359 WESGENSELKVAYVSMEDAKATYELKEFLPMETQLSRVQGLPMDVSSSGNLYEWYLL 358
OY 421 VSDTLERBCKKNYDVAPIVGYKFKDPGFIPISTLIGELITWQDIIKKMKATDPIEKK 480
DB 421 VSDTLERBCKKNYDVAPIVGYKFKDPGFIPISTLIGELITWQDIIKKMKATDPIEKK 480
OY 499 VSPDTLNEGCKKNYDIAQVGHKFCCKDIPGFIPISTLIGELITWQDIIKKMKATDPIEKK 478
DB 499 VSPDTLNEGCKKNYDIAQVGHKFCCKDIPGFIPISTLIGELITWQDIIKKMKATDPIEKK 478
OY 481 MDYRQAVKILANSYGYGAKAMWCKECAESTYANGRKYIELVWKELEBEKGFVYL 540
DB 481 MDYRQAVKILANSYGYGAKAMWCKECAESTYANGRKYIELVWKELEBEKGFVYL 540
OY 479 LLDYRQKAKRLANSFYGYGAKAMWCKECAESTYANGRKYIELVWKELEBEKGFVYL 538
DB 479 LLDYRQKAKRLANSFYGYGAKAMWCKECAESTYANGRKYIELVWKELEBEKGFVYL 538
OY 541 YADTDGFYATIPGEEKPETIKKKAKKEFLKYINSKLPGLLEVEGEFVLRGFFPAKKRYAVI 600
DB 541 YADTDGFYATIPGEEKPETIKKKAKKEFLKYINSKLPGLLEVEGEFVLRGFFPAKKRYAVI 600
OY 539 YIDTDGLVATITPGEGSEELKKALEFVKYINSKLPGLLEVEGEFVLRGFFPAKKRYAVI 598
DB 539 YIDTDGLVATITPGEGSEELKKALEFVKYINSKLPGLLEVEGEFVLRGFFPAKKRYAVI 598
OY 601 DEGRITTRGLEVVRDMSSEIAKETQAVLEAILKEDSVKAVEIVKQVVEIAKYQVPL 660
DB 601 DEGRITTRGLEVVRDMSSEIAKETQAVLEAILKEDSVKAVEIVKQVVEIAKYQVPL 660
OY 599 DEEGKVTTRGLEIVRDMSEIAKETQAVLEITLKHGVEEAVRIYKIOKLANYEIRP 658
DB 599 DEEGKVTTRGLEIVRDMSEIAKETQAVLEITLKHGVEEAVRIYKIOKLANYEIRP 658
OY 661 EKLVIHEOITKDSYKAGPHVAIAKRLAAGKIRPGTIISYILRSGKISDPAVL 720
DB 661 EKLVIHEOITKDSYKAGPHVAIAKRLAAGKIRPGTIISYILRSGKISDPAVL 720
OY 659 EKLAIVEOTTRPLHEKKAIPHAAKLAAGVKIKPGMIGVILYLRGDPISRAILTA 718
DB 659 EKLAIVEOTTRPLHEKKAIPHAAKLAAGVKIKPGMIGVILYLRGDPISRAILTA 718
OY 721 SEYDPKHKHYDDPYIENOVLPVLRILEAFGRKEDLYQSSQOYGLDAML 772
DB 721 SEYDPKHKHYDDPYIENOVLPVLRILEAFGRKEDLYQSSQOYGLDAML 772
OY 719 EEYDPKHKHYDAEYIENOVLPVLRILEGFGRKEDLYQKTRQVGLTSM 770
DB 719 EEYDPKHKHYDAEYIENOVLPVLRILEGFGRKEDLYQKTRQVGLTSM 770

```

RESULT 11

US-07-803-627A-1

Sequence 1, Application US/07803627A

Patent No. 5948663

GENERAL INFORMATION:

APPLICANT: MATHUR, Eric J.

TITLE OF INVENTION: Purified Thermostable Pyrococcus

TITLE OF INVENTION: furiosus DNA Polymerase I

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner, L.L.P.
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/803,627A
 FILING DATE: 02-DEC-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/776,552
 FILING DATE: 14-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,073
 FILING DATE: 19-FEB-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,568
 FILING DATE: 03-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Barker, M. Paul
 REGISTRATION NUMBER: 32,013
 REFERENCE/DOCKET NUMBER: 04121.0004-00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)408-4000
 TELEFAX: (202)408-4400
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 775 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-803-627A-1

Query Match 78.3%; Score 3151; DB 2; Length 775;
 Best Local Similarity 74.6%; Pred. No.2,4e-237;
 Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

1 MIPDTDTITDKGRPIIFPKKNEEFKIELDPHOPPIYALIKDSDAIDEIKATKGRBHG 60
 1 MILDVYITTEGGKRVILFKKNGKFKIENDRTFRPIYIALKRDSDIIEVKKITGERBHG 60

61 KIVRVVDVAVKKKKKLGDRDVEVMKLIPEHPDVPALRGKIREHNAVIDIYEYDIPAKRY 120
 61 KIVRIYDVEKVEKKFLKPIYMKLLEHPDVPITREKVEHNAVYDIFEYDIPAKRY 120

121 LIDKGLIPMGDEELKLMADIEFTFHEGDFEGKELIMISYADEEARVYTWKNIDLPY 180
 121 LIDKGLIPMGEEELKLIADIEFTFHEGDFEGKELIMISYADENAKVITWKNIDLPY 180

181 VDVVSNREMIKRFVQIVREKDPVLITVNGDNFDLPYLIRAKKLGITLLGDKKEHPE 240
 181 VEVVSSREMIKRFILIRKEDPPIITVNGDSDFPYLAKRAKLGITLLIGDGS--E 238

241 PKIHRMDSFAVEIKGRIFHDLFPVVRRTINLPYTLAEVAVLGKTKSGAEETAAI 300
 239 PKMGRIGDMVAVEYKGRIFHDLVHVRITINLPYTLAEVAVLFGKPKREVAVDAELAKA 298

301 WETESKAKLAQYSMEARATYELGKFEFFMEAEFLAKLIGOSVWDVSRSSGTNIVEWLL 360
 299 WESGENLEERAKYSMEARATYELGKFEFFMEAEFLAKLIGOSVWDVSRSSGTNIVEWLL 358

361 RVAYERNELPNKPDDEEYRRRLTTLGKGVKPERGMENTIYLDPRCLYPSIITYTHN 420
 359 RKAYERNEVAPNKPSEDEYRRRLTTLGKGVKPERGMENTIYLDPRCLYPSIITYTHN 418

421 VSPDTLREGCCKNYDVAPIVGKCKDPGFIPIISILGELITMROEIKKKMKATDPIIEKK 480
 419 VSPDTLNEGCKNDIANQVGHKCKDIPGFIPIISLGLHLEBRQIKIKMKMETQDPIEKI 478

481 MLDYRQRAVKLIANSYGYMGYPKARWYSKECAESVTAMGHHYIEMTIKEIEKFGFVYL 540
 479 LLDYRQRAVKLIANSFYGYMGYPKARWYSKECAESVTAMGHHYIEMTIKEIEKFGFVYL 538

541 YADPDGFYATIPGKEPEIKKAKKFLKYINSKLPGLLELEVEGYLFGFVAKKRVAVI 600
 539 YIDPDGFYATIPGGESEIEKKALEFVYINSKLPGLLELEVEGYLFGFVAKKRVAVI 598

601 DEGRITTRGLEVVRQMSSEIAKETQAKVLAAILKEDSVKAEVETVKQVVEIAKQVPL 660
 599 DEGRVITTRGLEVVRQMSSEIAKETQAKVLAAILKEDSVKAEVETVKQVVEIAKQVPL 658

661 EKLIVHEDITKDLSEYKAIGPHVAIAKRLAKGIVKRPCTIISYLVLGSKISDRVILL 720
 659 EKLIVHEDITRPLHEYKAGPHVAIAKRLAKGIVKRPCTIISYLVLGSKISDRVILL 718

721 SEYDPKHKHXPDPDYIENQVLAVALILAEAGYRREDLKYSKQVGLDAML 772
 719 EYDPKHKHXPDPDYIENQVLAVALILAEAGYRREDLKYSKQVGLDAML 770

RESULT 12
 US-08-688-649-37
 Sequence 37, Application US/08688649
 Patent No. 5827716

GENERAL INFORMATION:
 APPLICANT: MAMONE, JOSEPH A.
 TITLE OF INVENTION: MODIFIED POL-II TYPE DNA
 TITLE OF INVENTION: POLYMERASES
 NUMBER OF SEQUENCES: 38
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/688,649
 FILING DATE: Filed Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 220/281
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 776 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-688-649-37

APPLICATION NUMBER: US 07/686,340
FILING DATE: 17-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/626,057
FILING DATE: 11-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/513,994
FILING DATE: 26-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, GREGORY D.
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 927-5054
TELEFAX: (508) 927-1705
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1022 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
-08-271-364A-8

Query Match 62.2%; Score 2502.5; DB 1; Length 1022;
Best Local Similarity 62.9%; Pred. No. 1.2e-186;
Matches 534; Conservative 59; Mismatches 129; Indels 127; Gaps 17;

QY 1 MFDPDYITKDGKPIIRIKKKEGKFIELDPHOPYIYALKDSIDEIKAIKGERHG 60
DB 1 MLDIDYITKDGKPIIRIKKKEGKFIELDPHOPYIYALKDSIDEIKAIKGERHG 60
QY 61 KIVRVDAVKKKFLGDRDEVKKLIFEHPDVPALGKIREHNAVIDIYEDIPAKRY 120
DB 61 KIVRVDAVKKKFLGDRDEVKKLIFEHPDVPALGKIREHNAVIDIYEDIPAKRY 120
QY 121 LIDKGLIPMEGDEELKLMFDIETFEHEDGEGKELIMISYADEEARVITKNTIDLPY 180
DB 121 LIDKGLIPMEGDEELKLMFDIETFEHEDGEGKELIMISYADEEARVITKNTIDLPY 180
QY 181 VDVVSNREMIKRFVQIVREKDPVLITNGDNFDLPYLIRAKELGVTLLGRDKHEPE 240
DB 181 VDVVSNREMIKRFVQIVREKDPVLITNGDNFDLPYLIRAKELGVTLLGRDKHEPE 240
QY 241 PRIHRGDSFAVEIKGRIHFDLFPVVRRTINLPTYLEAVYEAVLGKTSKLGAEETIAI 300
DB 241 PRIHRGDSFAVEIKGRIHFDLFPVVRRTINLPTYLEAVYEAVLGKTSKLGAEETIAI 300
QY 301 WETEEEMKLAQYSMDARATYELGKEFFPMELAKLIGOSWDVSRSTGNLVEMVYL 360
DB 301 WETEEEMKLAQYSMDARATYELGKEFFPMELAKLIGOSWDVSRSTGNLVEMVYL 360
QY 361 RYAVENRELAPNKPDEEERRRRLRTYLGYYKEPERGLMENTYLDPRCLYPSIIVTNH 420
DB 361 RYAVENRELAPNKPDEEERRRRLRTYLGYYKEPERGLMENTYLDPRCLYPSIIVTNH 420
QY 421 VSPDTEREGCKNVDAPITGVYKFCCKDPFGFIPISIGELITMROETIKKMKATIDPIEK 480
DB 421 VSPDTEREGCKNVDAPITGVYKFCCKDPFGFIPISIGELITMROETIKKMKATIDPIEK 480
QY 481 MLDYRORAVKLANSTY-----GYMGYPK----- 504
DB 481 MLDYRORAVKLANSTY-----GYMGYPK----- 504
QY 505 -----ARWYKCEKAEK-----VTAMGRHYIEMTIKEIEKFGFV-----LYADTDGFY 548
DB 505 -----ARWYKCEKAEK-----VTAMGRHYIEMTIKEIEKFGFV-----LYADTDGFY 548
QY 541 VNNLFAFSEFNKKIKSEVKKVKKALIRHKKYKKAYPEIQLSGKKINTTAGHSLETVRNGEI 600
DB 541 VNNLFAFSEFNKKIKSEVKKVKKALIRHKKYKKAYPEIQLSGKKINTTAGHSLETVRNGEI 600
QY 549 ATIPGE-----KPEITIKKAKEFLKYINSKLPGLLELEYEGFYL-----RGFF- 591
DB 549 ATIPGE-----KPEITIKKAKEFLKYINSKLPGLLELEYEGFYL-----RGFF- 591
QY 601 KEVSGGCIKEGDLIVAPKKIKLNEKGVSTINIPELISDLSSEETADIVMTISAKGRKNFK 660
DB 601 KEVSGGCIKEGDLIVAPKKIKLNEKGVSTINIPELISDLSSEETADIVMTISAKGRKNFK 660

QY 592 --VAKRRIYVIDEGRIIT-----RGLVVRDWSSEIANETA-KVL 630
DB 661 GMLRTLRMMFGEENRRIIFTFNRYLFHLEKLGILKLPRGVEYV--DWERLKRYKQLYEKL 718
QY 631 EALIKEDSVKNAVEIKVDVEELAKYQVPLEKVIHEQITKDSKKAIQPHALAKRLA 690
DB 719 AGSVKYNKMKREYLVNFNEIKDFISY-FP-----QKELEWK-IGTLNGFPTNCI 766
QY 691 AKGIRKPGTIIISYIVLRG-----SGKISDRVILLSEYDPKHKHYDPDYIENQVLP 742
DB 767 LK-VDEDFGKLGLGYVSEGYAGAKKKTGIGS-----YSVKLYNEDPN-----VLE 811
QY 743 AVLRITLFAF 751
DB 812 SKMKNVAEKF 820

Search completed: May 28, 2002, 14:17:33
Job time: 31 sec